Tijdschrift voor Economie en Management
Vol. XLIX, 4, 2004

CAPM Tests and Alternative Factor Portfolio Composition: Getting the Alpha's Right

by L. De Moor and P. Sercu

Lieven De Moor
KULeuven, Department of Applied Economics and FWO-Vlaanderen, Leuven.

Piet Secru
KULeuven, Department of Applied Economics, Graduate School of Business Studies, Leuven.

ABSTRACT

We show that the results of a CAPM test are quite sensitive to the details of the test design. Especially crucial are the aspects related to the weight one gives to small, low-reputation stocks when constructing both the factor portfolios and the test or style portfolios whose returns are to be explained. To fit our observed returns we need to redesign the size and distress factor portfolios into two factor portfolios each, one for extremely small or distressed stocks relative to non-extreme stocks, and one for moderately small or distressed stocks versus larger or growth compamies. This alternative model does a better job in pricing stocks, both in the US and internationally, than the standard four-factor CAPM model with factor portfolios designed following Fama and French ((1992), (1993), (1995), (1996a), (1996b), (1998), (2000)), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999).

I. INTRODUCTION

The empirical anomalies that emerged from CAPM tests, such as the size, distress and momentum effects (Banz (1981); Stattman (1980) and Rosenberg, Reid and Lanstein (1985); Jegadeesh and Titman (1993)), have quickly been incorporated into generalized asset pricing models. Empirical work by *e.g.* Fama and French ((1992), (1993), (1995), (1996a), (1996b), (1998), (2000)), Carhart (1997) or Rouwenhorst (1999) reveals that these additional factor portfolios significantly improve the model's ability to capture the cross-sectional variation of stock returns, both within the US and internationally. The purpose of our paper is to extend these tests to a data set that has a uniquely wide coverage both across the size spectrum and across countries. We find that both in the US and internationally the ten percent smallest stocks do not fit the standard models, and there generally appear to be non-linearities missed by the standard three-factor FF model. In general, we need more than one return differential – that is, we need more than two portfolios – to capture the relationship between return and exposure to size or distress across the entire spectrum. The resulting generalized model provides a risk-return relation that outperforms national and global one-, two- or four-factor CAPMs (with market, size, distress, and momentum portfolios), the international CAPM with its exchange-factor portfolios, and the nested version of the international CAPM and the global four-factor model.

The structure of the paper is as follows. Section II focuses on the US market, the subject of most of the extant research. Our starting point is a replication of the Fama and French ((1993), (1996)) tests on a data set that potentially includes more – and, notably, smaller – stocks than those provided in the standard sources. When following the Fama and French (henceforth FF) procedure as closely as possible *re* data coverage we do find similar results as the original study, despite the different period (1980-2000 rather than 1963-1993), as shoiwn in Section II.A. In Section II.B. we then gradually modify the procedure, and notably increase the data coverage and the room given to small stocks. The result is large positive alphas for the lowest size decile and smile/smirk patterns across the board. Thus, we may need to add not just momentum but also an extra small-firm and distress factor to the original FF trio (market, SMB and HML). Our tests (Section III) of this candidate factor specification in the US market reveal that the extended model does explain various style portfolio returns, whether stratified across one or two styles and whether separated from the factor-portfolio data or not. In Section IV, then, we venture beyond the US borders and successfully test our extended asset pricing model against competing models, using various style portfolios (size, book-to-market, and momentum) as well as industry and country index returns. Section V concludes.

II. THE FAMA-FRENCH MODEL AND THE SMALL-FIRM ANOMALY REVISITED

Our first test design tries to be as close as possible to Fama and French (1993). We then test the robustness to modified designs. Some modifications are inspired by data availability outside the US, but the main change is the increased room for smaller stocks.

A. *Data*

Most studies look at the CRSP stocks (NYSE, Amex and NASDAQ) or, for international studies, Datastream stocks to the extent that book values are available. This data restriction eliminates primarily small stocks, and this could be problematic since even in standard data bases the low-cap end of the spectrum displays anomalies. In addition, the Fama-French tests (and many thereafter) discard data on financial corporations. Lastly, the standard Datastream data base is the "market list" which contains only stocks that are alive at the time of downloading, implying survivorship bias. We therefore include all NYSE, Amex, NASDAQ and NASDAQ Small-Cap stocks from Datastream's "research lists", after careful cleaning-up and filtering of these data. This US data set is part of the larger one, covering 264 months (1980-2000) and 39 countries, and described in a separate appendix available on request.

B. *Fama-French replication*

To set the stage we replicate the Fama and French (1993) test on our database, initially using the same termination date as they do in their 1996a study, end 1993. We start with an explicitly review of the main steps in their procedure since we will return to many of these in our robustness checks below.

Monthly dollar returns on 25 portfolios of size-and-distress-sorted stocks are regressed on three factor portfolios: the market portfolio, the size factor and the distress factor. To mimic Fama and French (1993) as closely as possible we use, in this first test, all stocks for which Datastream provides market caps and (positive) book values. At the end of June of each year, all these stocks are allocated to either of two groups (small or big, denoted S or B) depending on whether their early-June market cap is below or above the median market equity for NYSE stocks.[1] All stocks are also allocated, via an independent second sort, into one of three book-to-market (B/M) equity groups (low, medium, or high, denoted L, M, or H); the watershed values are the 30th and 70th percentile values of B/M-ranked NYSE stocks.

For the purpose of constructing the factors, six size·B/M portfolios are then defined as the six intersections of the two size groups and the three B/M groups. These six intersections are labeled S/L, S/M, S/H, B/L, B/M, and B/H. Value-weighted monthly returns on the six portfolios are calculated from July till June next year. For each month, the size factor SMB is computed as the difference between the returns on small stocks (the average of the returns on the three small-stock portfolios, S/L, S/M and S/H) and big stocks (the average returns on the three big-stock portfolios, B/L, B/M and B/H). The distress factor HML is the difference between, on the one hand, the average of the returns on the two high B/M portfolios (S/H and B/H) and, on the other, the average of the returns on the two low B/M portfolios (S/L and B/L). Note that the returns are value-weighted within each of the six size-B/M portfolios, while for the calculation of SMB and HML, equally weighted averages are taken across the three S/. or B/. portfolios.

For the purpose of generating test portfolios (that is, portfolios whose returns need to be explained by the factors), 25 size-B/M portfolios are formed following the same procedure as for the six size-B/M portfolios underlying SMB and HML, except that quintile breakpoints for size and B/M for NYSE stocks are used to allocate all stocks to the portfolios rather than the median or the 30th and 70th percentile values. Negative-B/M firms are discarded when calculating the breakpoints or forming size-B/M test portfolios.

The regression equation is:

$$R_i - r_f = a_i + b_i\,(R_m - R_f) + g_i SMB - d_i HML + e_i \tag{1}$$

Generally monthly portfolio returns do not exhibit significant autocorrelation. This was confirmed by the insignificant Durbin-Watson coefficients of each equation. But the returns do exhibit conditional hetero-skedasticity over time. Following Fama-French we initially ignore this, but towards the end we do shift to a hetero-skedasticity-consistent covariance matrix produced by the GMM version of OLS/SUR.[2] It appears that the GMM covariance matrix leads to fewer significant alphas than the OLS ones, but the difference is never very pronounced.

TABLE 1
Alpha estimates of Fama and French (1996a)

Size	Book-to-market				
	Low	2	3	4	High
Small	*–0.45*	*–0.16*	–0.05	0.04	0.02
2	–0.07	–0.04	0.09	0.07	0.03
3	–0.08	0.04	0.00	0.06	0.07
4	0.14	*–0.19*	–0.06	0.02	0.06
Big	*0.20*	–0.04	–0.10	–0.08	–0.14

Italic signals significance at a 5% level using the OLS standard error.

TABLE 2
Alpha estimates of Fama and French replication

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.02	–0.25	0.14	0.18	*0.26*
2	*–0.47*	–0.08	*0.30*	0.14	0.06
3	*–0.33*	0.02	–0.06	0.02	0.02
4	–0.09	–0.01	–0.10	–0.07	0.00
Big	*0.19*	–0.01	0.03	–0.07	*–0.39*

Italic signals significance at a 5% level using the SUR standard error.

For the reader’s convenience, Table 1 reproduces the alphas obtained in the original Fama and French (1996a) study. Table 2 then shows our own alpha estimates from the new data. The underpricing (or return shortfall) that, in FF (1996a), occurred for the small, growth stocks seems to have shifted up one class, into the second size quintile, possibly because part of our first size quintile is missing in the

FF database. In addition, the return anomalies for the distress stocks (the rightmost column) have become more pronounced, both algebraically and statistically. There may also be evidence of what looks like interactions: the extreme size-distress combinations show most mispricing, with the corner cases on the main diagonal being overpriced and those on the secondary diagonal underpriced. Still, the differences are not massive.

In the next subsections we verify whether these results are robust to minor modifications in the research design. We then gradually extend the coverage and the weight given to small stocks. When we do this at the factor-portfolio side, the fit improves, suggesting that the broadened factors do better. But a similar extension of the coverage on the left-hand-side (the test portfolios) worsens the fit, leaving us with an inadequate model.

C. *The impact of tangential design variations relative to FF*

Table 3 lists some minor differences between our tests and the original FF design. Many of these will be modified so that their impact can be tested. Table 4 to Table 13 demonstrates the evolution of the three-factor-model alphas, when in each step an extra design element is altered. Our starting point is Table 2, repeated for convenience as Table 4. Each change is maintained in subsequent tests – that is, changes are cumulative – with one exception that will be noted when it comes up.

We start with the time period. Table 4, Table 5 and Table 6 only differ regarding the years of data, with Table 4 showing the pre-1994 alphas

TABLE 3

Design differences with Fama and French (1996a)

Fama and French (1996a)	De Moor and Sercu
Financial firms excluded[3]	Financial firms included
US T-bill rate from CRSP, end of month	US T-bill rate from IMF, monthly average
Period: 7/63-12/93	Period: 7/80-12/93
Value-weighted returns: CRSP	Value-weighted returns: DataStream
No IPO’s; at least 24 months of data	IPO’s allowed
Book values from Compustat	Book values from DataStream

TABLE 4

Fama and French replication: data from 1980-1993

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.02	–0.25	0.14	0.18	*0.26*
Small	–0.02	–0.25	0.14	0.18	*0.26*
2	*–0.47*	–0.08	*0.30*	0.14	0.06
3	*–0.33*	0.02	–0.06	0.02	0.02
4	–0.09	–0.01	–0.10	–0.07	0.00
Big	*0.19*	–0.01	0.03	–0.07	*–0.39*

TABLE 5

Fama and French replication: data from 1994-2000

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.12	0.24	0.12	*0.37*	0.25
2	*–0.50*	–0.27	–0.24	–0.23	0.12
3	–0.33	*–0.63*	–0.30	–0.37	–0.15
4	–0.35	–0.44	*–0.55*	–0.28	0.19
Big	*0.37*	–0.25	*–0.56*	*–0.48*	0.11

TABLE 6

Fama and French replication: data from 1980-2000

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.03	–0.10	0.13	*0.24*	*0.27*
2	*–0.48*	–0.15	0.14	0.01	0.07
3	*–0.32*	–0.18	–0.15	–0.10	–0.02
4	–0.18	–0.16	*–0.27*	–0.16	0.05
Big	*0.26*	–0.10	–0.18	–0.20	–0.21

Italic signals significance at a 5% level using the SUR standard error. GMM, used in Table 13, takes into account cross-equation correlation and intertemporal hetero-scedasticity. “Narrow-based” refers to observations with both market- and book-value data. “Broad-based” data use all data whenever possible even if book value is missing. “NYSE” or “all” refers to the list – broad or narrow – from which the required deciles are computed.

(the overlap with FF (1996a)), Table 5 displsying the post-1993 results, and Table 6 the alphas for the full sample. Apparently the chosen time period in the design of a CAPM test does not influence the results much. The same seems to hold for the choice of the risk-free rate and the market index. Specifically, when going from Table 6 to Table 7 the risk-free rate becomes the US discount rate instead of the US T-bill rate,[4] and the market return is Datastream's US market return, not the value-weighted return on all stocks in the size-distress portfolios plus the negative-book-value equities as in Fama and French ((1993), (1996a)). Again we cannot detect much difference in the alphas.

In the data underlying Table 8 the compositions of both the test and the factor portfolios are updated every month instead of yearly. Compared with the result of Table 7, it is clear that the three-factor model no longer seems to do a good job in pricing the 25 unmanaged size-distress portfolios if portfolios are updated more frequently. As there is no intrinsic reason why this should be so, we keep using monthly updating in the tests below, as an anomaly or at least an issue of robustness that should be resolved.

D. *Increasing the coverage and weight for small firms in FF*

The next three design features whose impact should be verified all have to do with the weight, or lack thereof, given to small stocks. First, the FF procedure is to discard stocks for which either book value or market cap is missing, a restriction that tends to eliminate mostly small companies. Thus, the standard SMB and HML may overlook part of a small-firm effect. Simultaneously, any such deficiency in the factors may never show up because the companies most affected by the potentially missed factor are missing on the left-hand side too. Second, in FF, the assignment of stocks to factor portfolios or test portfolios is based on NYSE percentile values even though the data base also includes Amex and NASDAQ stocks. This results in size groups with more firms in the smaller categories and, likewise, distress groups with more stocks in the growth or low book-to-market category. The third design feature in FF that may underplay any small-firm effect is value weighting. While the portfolio-theory logic underlying the CAPM dictates value weights as far as the market portfolio is concerned, there is no such theoretical basis for the size and distress factors. One drawback of value weighting is that the S factor portfolio, even though it contains all below-median stocks, is dominated by the comparatively larger ones,

TABLE 7

Using the Datastream market return and the USD discount rate for R_m and R_f

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.04	–0.12	0.12	0.22	*0.24*
2	*–0.50*	–0.17	0.13	0.00	0.05
3	*–0.34*	–0.20	–0.17	–0.12	0.05
4	–0.20	–0.18	*–0.30*	–0.18	0.03
Big	*0.24*	–0.13	–0.21	–0.22	–0.22

TABLE 8

Monthly updating of the size- and distress factor portfolios and test portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.25	*–0.39*	–0.21	0.11	*0.39*
2	*–0.41*	–0.15	*–0.41*	–0.16	*0.18*
3	*–0.29*	*–0.41*	*–0.50*	–0.20	–0.04
4	–0.04	*–0.30*	*–0.67*	*–0.35*	0.01
Big	*0.28*	–0.21	*–0.29*	*–0.37*	–0.04

Italic signals significance at a 5% level using the SUR standard error. GMM, used in Table 13, takes into account cross-equation correlation and intertemporal hetero-scedasticity. "Narrow-based" refers to observations with both market- and book-value data. "Broad-based" data use all data whenever possible even if book value is missing. "NYSE" or "all" refers to the list – broad or narrow – from which the required deciles are computed.

those close to the median size.[5] Since, in addition, the median is the NYSE one, the value-weighted S portfolio really is more of a mid-cap portfolio than the small-cap one like its name would suggest. For these reasons we experiment with more equal-sized portfolios, equal weighting, and stocks with missing book data, fully realizing that this may carry its drawbacks: small firms may suffer from excess noise because of thinner markets and patchier attention from analysts.

We start with the weighting scheme, introducing equal weights in turn on the test and factor sides. (Thus, for this once the changes between Table 9a and Table 9b are not cumulative.) For *test* portfolios, if equal weighting produces more rejections of the model, the change in the design should be maintained in the sense that the new test apparently has more power. If equally-weighted *factor* portfolios, in contrast, lead to more rejections, this change is not to be maintained since it means the new factor is mis-specified. Going from Table 8 to Table 9a or 9b, we see that either change in itself has but a small effect on the number of rejections (which goes up from 14 to 15 in each case). Introducing equal weighting on both sides simultaneously, as in Table 9, adds one more rejection; in addition, the unexplained returns are also economically larger.

We now show that these problems diminish when we also extend the size coverage of the factor portfolios, and worsen again when we do the same on the test-portfolio side. In Tables 10 to 13, the factor portfolios are built from all stocks, not just those with both market- and book-value data. Specifically, the size factor is now computed as the difference between the equally weighted average return for all stocks above the median versus the average for all below-median stocks, whether they provide book-value information or not; similarly, the distress factor is the difference of the equally-weighted returns on portfolios containing the firms that rank below the 30^{th} or above the 70^{th} percentile *re* B/M. These percentiles are, for the time being, still based on NYSE stocks with full data.

In this new test, the coverage for distress is the same as before, since market values are almost never missing. For the size variable, in contrast, the number of stocks goes up by over 60 percent on average. Indeed, on average, 40% of Datastream's NYSE stocks have no accounting data. This average hides a strong time trend: in the early 1980s, two Datastream records out of three lacked book values, but this ratio is down to one out of ten by 2000. Similar numbers hold for non-NYSE stocks. Since the total number of stocks in the 1980s is lower too, we can expect a substantial improvement of the quality of the size portfolios in the beginning of the sample period if we drop the FF data requirement. In addition, the missing firms are predominantly small: when dropping the data filter, the mean market cap falls by about 50 percent on average – in fact, by 80% in the early years, 10% in the most recent ones.

Comparing the alphas of Table 10 with those of the table before shows that the broadened factor portfolios are more capable of pricing unmanaged size-distress portfolios. The number of rejections drops

TABLE 9A

Equally-weighted test portfolios, value-weighted factor portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	–0.06	–0.19	–0.08	0.20	*0.72*
2	*–0.44*	–0.14	*–0.38*	–0.12	*0.22*
3	*–0.23*	*–0.39*	*–0.50*	*–0.22*	0.02
4	–0.05	*–0.30*	*–0.66*	*–0.40*	0.06
Big	0.16	*–0.43*	*–0.49*	*–0.53*	0.07

TABLE 9B

Equally-weighted factor portfolios, value-weighted test portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	*–0.48*	*–0.70*	*–0.53*	*–0.29*	–0.12
2	*–0.37*	–0.29	*–0.58*	*–0.37*	–0.09
3	–0.15	*–0.48*	*–0.62*	*–0.32*	–0.27
4	0.19	–0.26	*–0.72*	*–0.42*	–0.19
Big	*0.32*	–0.20	*–0.30*	*–0.38*	–0.14

TABLE 9

Equally-weighted factor and test portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	*–0.35*	*–0.54*	*–0.48*	*–0.27*	0.14
2	*–0.39*	–0.26	*–0.56*	*–0.33*	–0.05
3	–0.09	*–0.46*	*–0.62*	*–0.33*	–0.19
4	0.18	–0.23	*–0.70*	*–0.46*	–0.12
Big	*0.32*	*–0.39*	*–0.47*	*–0.56*	–0.04

Italic signals significance at a 5% level using the SUR standard error. GMM, used in Table 13, takes into account cross-equation correlation and intertemporal hetero-scedasticity. "Narrow-based" refers to observations with both market- and book-value data. "Broad-based" data use all data whenever possible even if book value is missing. "NYSE" or "all" refers to the list – broad or narrow – from which the required deciles are computed.

markedly, from 16 to 9. This strongly suggests that the new factor specification is a step in the right direction: the FF factors, by restricting the coverage to stocks with both a known market value and a known book-to-market value, miss too many of the smaller firms. Computing the 30^{th}, 50^{th}, and 70^{th} percentile values from all NYSE stocks even if not both values are available, as is done for the alphas in Table 11, further decreases the number of rejections from 9 to 7. Again, giving more room to the small stocks in the factor portfolios improves the picture.

Similar changes can also be implemented on the test-portfolio side. The most powerful results (in the sense of providing the highest number of rejections) were obtained as follows. We keep the earlier 25 pure-intersection portfolios as the starting basis of the new test portfolios. The additional stocks, those with just size information, are sorted into the five size buckets, and from there are transferred to one of the 25 old intersection portfolios, taking care to stay within the same size bracket but randomizing across B/M category. This procedure shrinks the dispersion across distress classes, but everything else being the same, also reduces the noise in the portfolio returns.[6] Thus, whether on balance power improves or not is an empirical matter. The outcome, in Table 12, is a dramatic increase in the number of rejections, which doubles to 14.

TABLE 10

Broad-based factor portfolios; narrow-based breakpoints (NYSE) and test-portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	0.18	–0.06	–0.11	0.04	*0.40*
2	0.04	0.06	–0.29	–0.13	0.16
3	0.21	–0.24	*–0.45*	–0.23	–0.04
4	*0.32*	–0.13	*–0.64*	*–0.40*	–0.01
Big	*0.31*	*–0.39*	*–0.49*	*–0.56*	–0.01

Italic signals significance at a 5% level using the SUR standard error. “Narrow-based” refers to observations with both market- and book-value data. “Broad-based” data use all data whenever possible even if book value is missing. “NYSE” or “all” refers to the list – broad or narrow – from which the required deciles are computed.

TABLE 11
broad-based factor portfolios and breakpoints (NYSE); narrow-based test-portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	0.67	0.13	–0.13	0.01	*0.59*
2	0.15	–0.24	–0.22	–0.07	0.06
3	0.08	–0.19	*–0.40*	–0.16	0.13
4	0.22	–0.28	*–0.51*	–0.23	0.09
Big	*0.28*	*–0.31*	*–0.51*	*–0.46*	–0.04

TABLE 12
Broad-based factor portfolios, breakpoints (NYSE), and test-portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	0.13	0.03	–0.15	–0.15	*0.28*
2	–0.25	*–0.44*	*–0.55*	*–0.53*	–0.31
3	–0.18	*–0.34*	*–0.46*	*–0.34*	–0.29
4	0.13	*–0.39*	*–0.56*	*–0.39*	–0.08
Big	*0.23*	*–0.31*	*–0.40*	*–0.40*	0.02

TABLE 13
Broad-based factor portfolios, breakpoints (all), and test-portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	*0.40*	*0.42*	*0.38*	–0.03	*0.62*
2	*–0.47*	*–0.40*	*–0.55*	*–0.62*	*–0.48*
3	–0.32	–0.15	*–0.51*	*–0.37*	–0.15
4	0.20	–0.17	–0.32	–0.23	–0.27
Big	*0.30*	*–0.25*	*–0.41*	*–0.47*	0.03

Italic signals significance at a 5% level using the SUR (GMM) standard error. GMM, used in Table 13, takes into account cross-equation correlation and intertemporal hetero-scedasticity. “Narrow-based” refers to observations with both market- and book-value data. “Broad-based” data use all data whenever possible even if book value is missing. “NYSE” or “all” refers to the list – broad or narrow – from which the required deciles are computed.

In a last design change, we base also the breakpoint values on the quintile values from the entire data set, not just the NYSE ones.[7] Again, the size coverage of the portfolios widens because the number of assets per size or B/M group is now equal across groups rather than very much bunched together at the small-cap or high-growth end. The effect of the new way of defining the buckets becomes stronger over time, this time: in 1980 the non-NYSE list in Datastream represents just 13% of the total, but that percentage rises to over 70% in 2000. Non-NYSE firms in Datastream had a mean market cap of less than one-fourth of the typical Big-Board listee in 1980, and about 45% in 2000. Thus, as expected, computing the quintiles from the all-stock list brings about drastically lower quintile values for especially the first quintiles. The result of this procedure, shown in Table 13, is not so much a better fit – at 13, the number of rejection remains virtually unaffected – as a shift in the rejections, which now occur mostly in the lower-cap of the table.

The large number of significant abnormal returns is not the only anomalous result. In addition, the typical rejected alpha is about 0.5%/month or more, which is worse than the kind of numbers FF obtain. Lastly, there are manifest patterns in the alphas. First, within each and every row there is a smile pattern in the alphas. Second, within each column there is a smirk pattern, with the small-firm quintile always providing a strongly positive excess returns, the second quintile a strongly negative one, followed by gradually improving returns for higher-size quintiles. It is, we think, fair to say that the three-factor model does not span our returns and that size seems to be part of the problem.

III. IDENTIFYING THE MISSING FACTORS: US DATA

In this section we propose a generalized FF model that takes care of most of the anomalies we just noted. In view of the momentum-related anomalies that came to light after the publication of FF (1993), momentum is added into the analysis throughout Section III. Following the Rouwenhorst (1999) version of Jegadeesh and Titman (1993), stocks are ranked on the basis of the return realized in the months t–7 to t–2. (Month t–1 is omitted to eliminate the common bid-ask-bounce effect that would otherwise have affected both the past performance and the subsequent return.) All available data are used, whether book value is available or not. The momentum factor is the equally-weighted return, for month *t*, on the 30% best winners minus the 30% worst losers. Contrary to the size- and distress-portfolios, momentum-portfolios have a holding period of not one month but six, as in Rouwenhorst (1999) or Jegadeesh and Titman (1993). Again following these authors, we compute the monthly average return across the six ongoing momentum strategies, each started one month apart, to handle the issue of overlapping observations. Like the size- and distress portfolios, the momentum portfolios are updated monthly and are equally weighted. When we make momentum test portfolios we use deciles or quintiles rather than the 30th or 70th percentiles.

We start, in Section III.A., with a look at mean returns on decile portfolios, one set per risk dimension. Even though the sorting is one-dimensional and the returns are not risk-adjusted, we find back the smiles and smirks we obvserved in the alphas of the previous section. In passing, we also provide evidence that the size coverage is the one most influential extension of the data set, and that adding a momentum factor does not solve the problems. In Section III.B. we look at risk-adjusted returns from three sets of two-dimensionally sorted portfolios, one set per pair of risk dimensions, and we still find the same nonlinearities. All this suggests that it may be hard to fit, say, the effect of size on return via one single factor, that is, one return. A closer scrutiny of one-dimensional decile portfolios provides ideas on how to define the additional factors (Section III.C). The resulting expanded model in its full version is then successfully tested against the standard models. The obvious risk, in this approach, is that we might be over-fitting a specific data set; however, bear in mind that the resulting model is tested also on international data (Section IV), where it appears to hold well, too.

A. *One-dimensionally sorted portfolios: the role of size revisited*

In this section we look at returns from decile portfolios of stocks sorted along one dimension at the time. We also provide a robustness check for our finding, thus far, that size coverage is the main reason why FF does not fit our model. Lastly, we seize the opportunity to compare the size bias also to another bugbear of Datastream, survivorship bias.

Figures 1 to 3 show average returns for ten decile portfolios sorted by size, distress (B/M) and momentum, respectively. For further reference we note three things. First, the main size effect is found in the first and to some extent also the second decile, which provide unusually large returns. In deciles 3-10, in contrast, there is a weak premium for larger sizes. Second, the distress effect is more monotone positive, but S-shaped rather than linear. There is a mild return-shortfall effect in deciles 1 and 2 (growth firms earning moderately lower returns), which then flattens out; and as of decile 7, "value" firms earn increasingly higher premia. Third, an S-effect is also present in the momentum factor, with strong losers going on earning clearly lower returns, strong winners continuing their upward trend, and flat returns for a wide midrange (deciles 4-8). Common to the three schedules is the nonlinearity. These patterns raise the possibility that the tradition of capturing the size factor (or distress or momentum factor) by just one number, the difference between a "hi" and a "lo" portfolio return, may be too simplistic. We return to this later on.

We next test the robustness of the previous section's finding *re* the importance of the size coverage in the data. In Section, the expansion of the data was done in line with Datastream's gradually extending coverage. In practice this means that in the early 80s data, the number of stocks added next to those meeting the original FF criteria was small while it became quite large in the early 2000s. To make sure that we are picking up a pure size effect and not some interaction between size and time, we now compare the full data set with an alternative size-biased sample where the rate of data rejection is more constant over time. We also compare the importance of this size-coverage effect to a survivorship effect that is present in some studies.

The standard stock list, in most Datastream-based research, is the country's "market list". This list suffers from two biases: it omits small stocks (as it climbs down the size list until 80 or 90% of the market cap has been picked up), and it consists of just stocks that exist in the year of downloading. To be able to estimate the relative influence of either bias we extract two non-random samples out of our full US set. The first set is survivorship-biased: it altogether excludes any stock that disappeared at any time from the full US-database during the period. For the size-biased data set, in contrast, at the end of each year we eliminate the 20% smallest stocks from the full US-database for one year. This second data set is free of survivorship-bias as the 80% largest stocks still can be delisted during the year.

Figures 4 to 6 plot the deciles' average returns for the survivorship-biased sample. Comparing to the full-sample graphs just above, it

FIGURE 1

Average returns for size-sorted deciles (no bias)

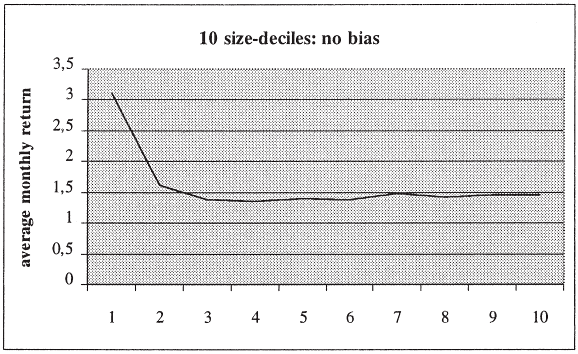


FIGURE 2

Average returns for distress-sorted deciles (no bias)

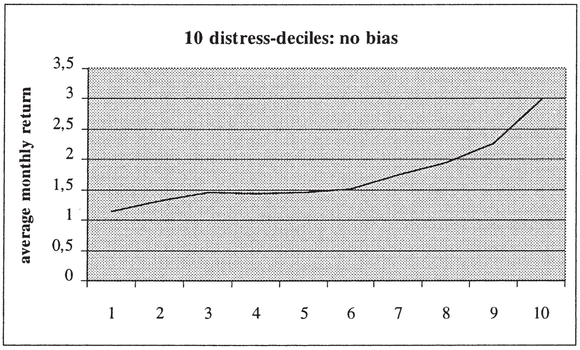


FIGURE 3

Average returns for momentum-sorted deciles (no bias)

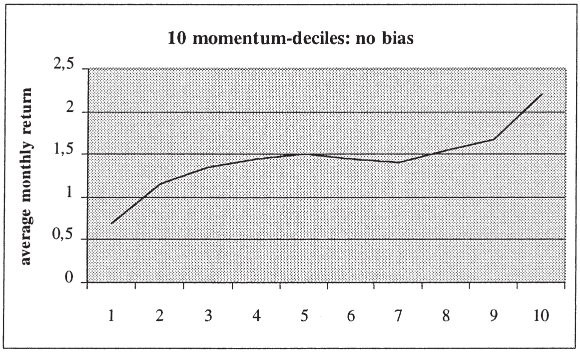


FIGURE 4

Average returns for size-sorted deciles (survivorship-bias)

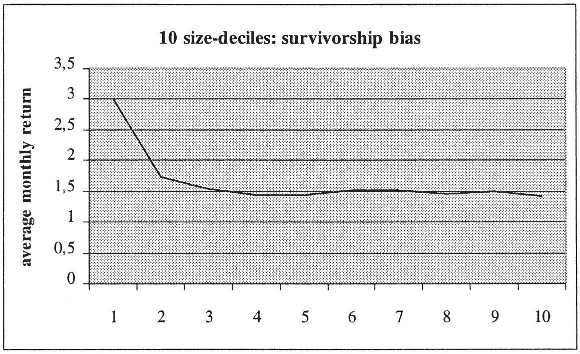


FIGURE 5

Average returns for distress-sorted deciles (survivorship-bias)

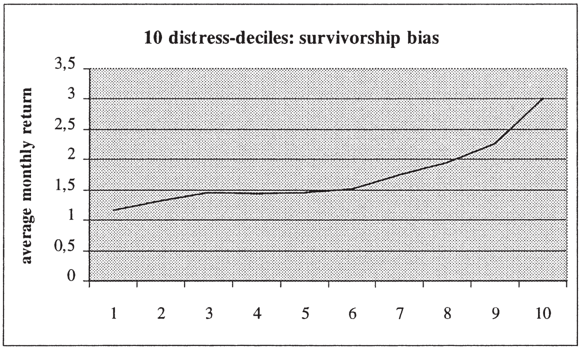


FIGURE 6

Average returns for momentum-sorted deciles (survivorship-bias)

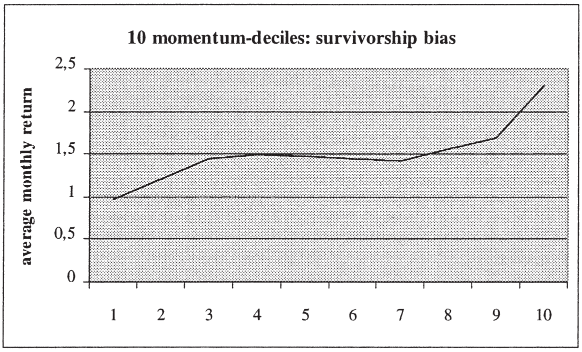


FIGURE 7

Average returns for size-sorted deciles (size-bias)

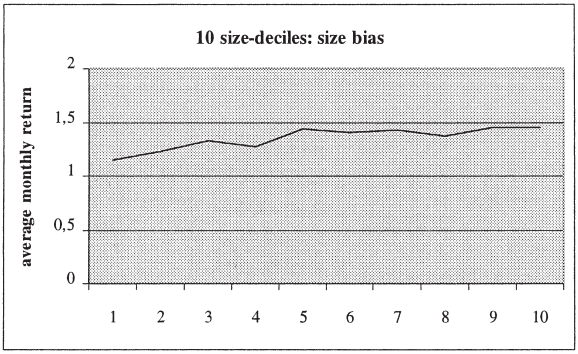


FIGURE 8

Average returns for distress-sorted deciles (size-bias)

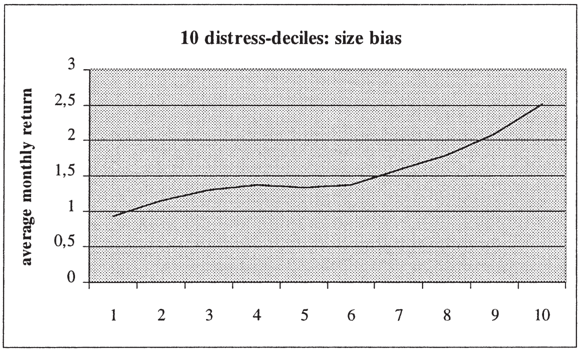
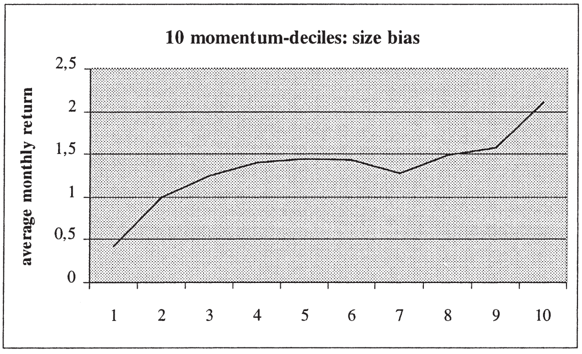


FIGURE 9

Average returns for momentum-sorted deciles (size-bias)



becomes obvious that survivorship bias does not have any substantial influence on the average monthly dollar return of ten size-based portfolios. The size-anomaly, notably, does not seem to be influenced by survival at all. But in the size-biased sample the positive size effect for the smallest stocks and the negative size effect for the larger stocks disappear completely when the database is size biased. The other return patterns seem unaffected by either size or survival-based filters. In the case of distress, one reason may be that small firms often have missing accounting data and, therefore, have never entered the distress portfolios in the first place.

Table 14 provides statistical evidence rather than graphs. It lists the mean return on the size, distress and momentum factor portfolios in each of the three data sets, along with *t*-ratios relative to H_0: $\mu = 0$. We see that the mean returns on all factor portfolios are significantly positive, except for the size-biased database where the SMB return differential is now insignificantly negative. Again, the conclusion is that the small-firm effect stems from, at most, the 20% smallest stocks. The distress- and momentum factor portfolios, in contrast, do not seem to be substantially influenced by survivorship- or size-bias. The average return of the momentum factor portfolio drops slightly when dead stocks are eliminated. Arguably, a burning-out process of dead stocks (successive months of negative returns) would strengthen the momentum factor portfolio. The average return of the momentum factor portfolio slightly rises when small stocks are eliminated. The interpretation is not clear: small stocks may weaken the momentum factor portfolio, or small stocks may be more likely to exhibit short term reversals, or larger stocks amay be more likely to exhibit short-term return persistence.[8]

Table 15 summarizes tests as to whether the mean or variance of the standard factor portfolios differ among the three versions of the US

TABLE 14

Monthly factor portfolio averages (and t-statistics) for three US-databases

	Free of bias	Survivorship biased	Size biased
SMB	0.59 (3.00)	0.64 (3.13)	–0.19 (–1.09)
HML	1.09 (5.60)	1.09 (5.59)	1.00 (5.33)
WML	0.74 (4.23)	0.63 (3.60)	0.83 (4.76)

TABLE 15
P-values for mean- and variance-F-difference tests

		Mean-F-test	Var-F-test	Correlation
SMB	FREE - SURV	0.86	0.57	0.97
	FREE - SIZE	*0.003*	0.07	0.92
	SURV - SIZE	*0.002*	*0.01*	0.92
HML	FREE - SURV	0.99	0.98	0.99
	FREE - SIZE	0.74	0.60	0.97
	SURV - SIZE	0.74	0.59	0.98
WML	FREE - SURV	0.68	0.83	0.97
	FREE - SIZE	0.68	0.90	0.97
	SURV - SIZE	0.68	0.92	0.95

database. We again conclude that only the size bias has a significant influence on the mean and variance of the size-factor portfolio. Note that even in that case the correlations between the various versions of the factor portfolio remain well above 90%.

To sum up: distress- and momentum factor portfolios are hardly affected by either survivorship or size bias. Also the size-factor portfolio does not seem to be influenced much by survivorship bias, as far as we can tell (which may not be very far). However, size bias does have a large effect on the size-factor portfolio. It again looks as if the bulk of the size effect in the US-market stems from, at most, the 20% smallest stocks. Stated differently, the results of the previous section are robust to the way the small stocks are eliminated from, or re-added to, the data base.

We also remember that, across deciles, returns seem to be evolving in a non-linear way, suggesting that one single return differential may be insufficient to summarize the size effect (or momentum or distress effect). True, this inference is indicative only. For one thing, in theory the stocks' sensitivities to the factors could be sufficiently non-linear in the decile's order i to pick up the apparent nonlinearity. Second, the sort is one-dimensional; in theory the omitted other risk factors could still be responsible for what here seems to be a non-linearity. Still, we obtained very similar conclusions from the alphas in the previous section, where exposures to factors were used rather than quintile membership and where two dimensions of non-market risk were considered simultaneously. In the next section we extend this two-dimensional analysis to include the momentum factor.

B. *Two-dimensionally sorted portfolios*

There are not enough data to work with a full $5 \times 5 \times 5$ classification: many cells in such a three-dimensional classification remain empty, and others have pityfully few members. As a second best we can still study, sequentially, three two-way classifications. Specifically, based on its end-of-period market value, book-to-market ratio and momentum every stock is assigned a membership of (*i*) one of 25 size·distress intersection portfolios; (*ii*) one of 25 size·momentum intersection portfolios and (*iii*) one of 25 distress·momentum intersection portfolios. The portfolios are updated monthly, weighted equally and consist of US stocks only, for the period 1980-2000.

The average returns for each set of 25 portfolios are shown graphically rather than as a set of numbers. For instance, the full piecewise-linear curve in Figure 10 connects the five mean returns, for each of the market-value buckets listed on the horizontal axis, of the stocks in the highest distress bracket (the highest B/M). the dotted curve indicates the returns for the lowest distrss firms, and so on. If all the schedules are roughly parallel, then the inference is that our earlier “marginal” patterns (from the one-dimensional sorts) are internally validated, and that the two-dimensional grid is the sum of two one-dimensional schedules. Non-parallel schedules, in contrast, would signal interactions on top of the additive main effects.

The results can be broadly summarized as follows. Firstly, differences between the mean returns are mainly driven by the extreme quintiles: the dotted and the full schedule, which always refer to the first and last quintiles, almost everywhere occupy the most extreme positions, while the three other curves are much closer and frequently intercross. This echoes our finding from the one-dimensional analysis. Second, the middle schedules tend to resemble each other in shape, while for the extreme ones (dotted and full), this often is far less the case. Thus, interactions, if any, seem to be active mainly in extreme portfolios. The third general pattern is that, like in the one-dimensional analysis, the schedules are often far from linear, meaning they may be poorly summarized by the returns and risks of just two portfolios. Lastly, there are some specific patterns. For instance, the

FIGURE 10

Interaction: size-distress

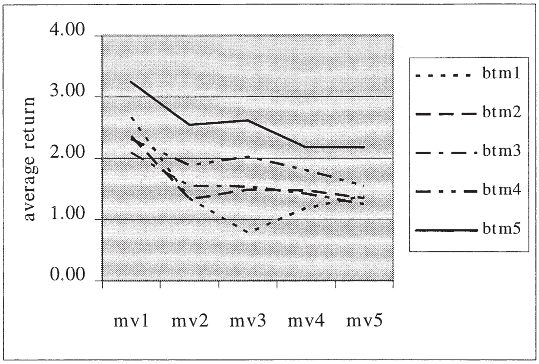


FIGURE 11

Interaction: size-momentum

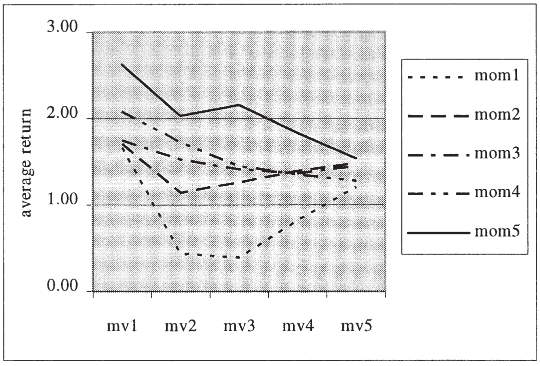


FIGURE 12

Interaction: distress-momentum

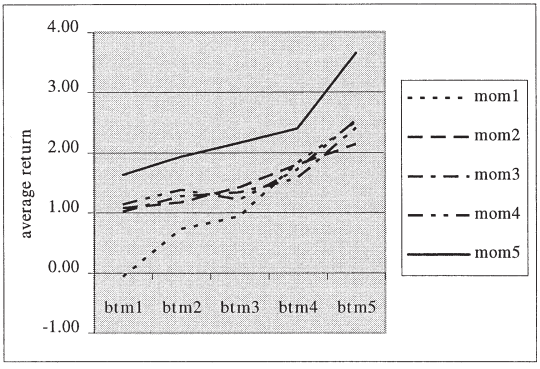


FIGURE 13

Interaction: size-distress

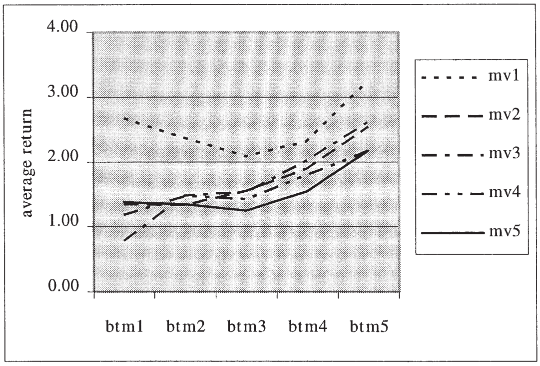


FIGURE 14

Interaction: size-momentum

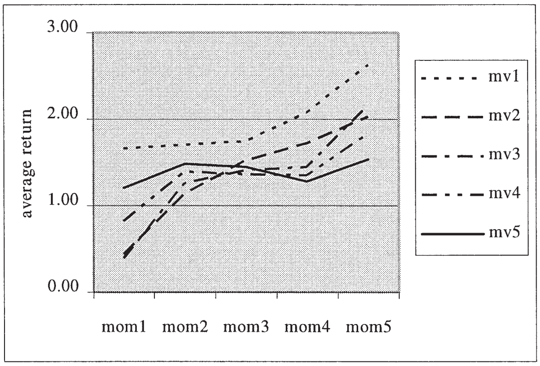
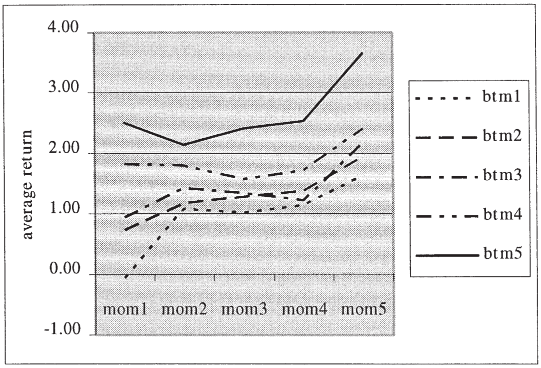


FIGURE 15

Interaction: distress-momentum



influence of the distress category and the momentum category is largest for midcap stocks; for the largest stocks, in contrast, momentum seems to have no influence on the expected stock return. Other findings are that the influence of size on the expected return is very small for “winner” half of schedule; and the influence of distress on the expected return is largest for loser stocks.

Again, the suggestions of nonlinearities are indicative only: decile membership is not the same as sensitivity to a factor, and the omitted third risk factor or the market beta could still be responsible for what here seems to be a non-linearity or an interaction in any of our two-dimensional grids. Still, the evidence seems sufficiently interesting to motivate a more detailed analysis, at the decile level and fully taking into account estimated exposure rather than decile or quintile membership. This is the topic of the next subsection, where we also try to construct factor portfolios so as to get the alphas of unmanaged funds as close to zero as possible.

C. *In search of optimal factor portfolios*

The conjecture behind the rest of the paper is that the apparent mispricing noted in Sections III.A and III.B stems from non-linear relations between factor exposure and return, and that this may be resolved by using, in every risk dimension, two return differentials rather than one to summarize the return-risk relationship.[9] We first explore this possibility by looking at one-dimensionally sorted decile portfolios in two-factor models where, next to the market, either size or momentum or distress is present. We then compare the performance of the full model, with its seven factors, to competing models. In Section IV we then test the approach largely out-of-sample, to wit on international data.

The construction of factor and test portfolio always proceeds as follows. Decile breakpoints are set using all stocks, including Amex and NASDAQ firms. Stocks with an unknown market value or book-to-market value are also used to set breakpoints and to calculate the other factor portfolios (that is, all breakpoints and factor portfolios are broad-based). And all portfolios are equally weighted and updated monthly. All regression test *t*-statistics are computed using a GMM specification that accounts for intertemporal hetero-skedasticity within each series beside, of course, cross-equation hetero-skedasticity and correlation.

1. A second size factor

The average monthly dollar returns of the size-deciles for the period 1980-2000 were already shown in Figure 1, which revealed a large average return for the first-decile (smallest) stocks and a slightly higher average return for the largest stocks compared to the middle deciles. This last observation is in line with Fama and French (1992), who find evidence that the size premium in the US has become weaker in recent years. In fact, for 1980-1990 they document a negative size premium.[10] Also Eun, Huang and Lai (2003) likewise find that, recently, the mean return is somewhat higher for large-cap funds than for small-cap funds in the US market[11]. But from Figure 1 and the existing literature, it seems that there still exists a strong small-firm effect for the first-decile stocks in the US that is missed by databases that are too selective. Only when we go beyond our first- and second-decile stocks we see an inverted small-firm effect in the US.

All this was about raw returns linked to decile membership, not returns risk-corrected via regression exposure coefficients. Table 16 exhibits the alphas' estimates and *t*-statistics for the ten size deciles for three different CAPM versions. The one-factor model is the basic CAPM model with market risk as the only source of cross-sectional variation. From Table 16 we see that the basic CAPM cannot price the smallest decile correctly; beta does not seem to be the only relevant exposure, in other words. The two-factor model shown next to the regular CAPM is the version with an extra size-factor portfolio, composed like in Fama and French (1993): it is the difference between the returns on the 50% biggest and lowest stocks. Table 16 shows that the standard two-factor CAPM does a bad job – worse than the single-factor CAPM, in fact – in pricing unmanaged size-portfolios: nine of ten alphas are significantly different from zero,[12] and the first-decile alpha has become even worse. The fact that so many alphas are affected, and to such an extent, demonstrates that most stocks do load on the FF size factor. Still, the result is clearly unsatisfactory.

On the basis of Figure 1 and the *t*-statistics of the one- and two-factor model, we now experiment with two long-short portfolios rather than one. The figure suggests two kinds of size-risk: (*i*) the regular size factor like in Fama and French (1993) which holds for all stocks but the smallest; and (*ii*) the risk inherent to the smallest stocks that cannot be accounted for by neither beta risk nor the regular FF size risk. Since FF already coined the label Small for their not-so-small stock portfolio, we reluctantly chose the label “micro stocks” for our new factor, even though by many countries’ standards these micro stocks are still quite sizable. We let *mSMB* denote the micro-stock risk factor, defined as the return on a zero-investment portfolio that is long the first decile and short deciles 2 to 10. We let *rSMB* denote the regular size risk factor, defined as a zero-investment portfolio that is long in stocks from deciles 2 and 3 and short stocks from deciles 6 to 9. Thus, in the alternative version the two-factor size model has become a three-factor model,

$$R_i - r_f = a_i + b_i\,(R_m - r_f) + g_i mSMB + d_i rSMB + e_i \qquad (2)$$

Table 16 shows that our model does a good job in pricing the ten size decile portfolios: none of the alphas is significantly different from zero. These conclusions remain valid when we use only NYSE stocks to calculate the decile breakpoints.

TABLE 16

Alphas’ estimates and t-statistics: size-deciles

	1-factor	2-factor	alternative
Small	*1.68 (6.07)*	*1.02 (9.84)*	–0.10 (–0.47)
2	0.08 (0.30)	*–0.55 (–4.76)*	–0.10 (–0.45)
3	–0.15 (–0.60)	*–0.70 (–5.56)*	–0.04 (–0.15)
4	–0.23 (–0.97)	*–0.72 (–5.20)*	–0.24 (–0.89)
5	–0.22 (–0.97)	*–0.63 (–4.31)*	–0.24 (–0.86)
6	–0.25 (–1.19)	*–0.55 (–3.51)*	–0.17 (–0.65)
7	–0.13 (–0.72)	*–0.35 (–2.30)*	0.03 (0.11)
8	–0.18 (–1.13)	*–0.34 (–2.56)*	–0.08 (–0.35)
9	–0.15 (–1.20)	*–0.23 (–2.08)*	–0.05 (–0.28)
Big	–0.09 (–1.76)	–0.10 (–1.91)	–0.01 (–0.12)
# Sig	1	9	0
Adj R^2	0,68	0,86	0,82
χ^2-test	0,00	0,00	0,44

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 17

Alphas' estimates and t-statistics: B/M-deciles

	1-factor	2-factor	alternative
Low	*–0.60 (–2,66)*	*0.54 (2,59)*	–0.10 (–0,52)
2	*–0.39 (–1,99)*	*0.47 (2,43)*	–0.10 (–0,54)
3	*–0.23* (–1,27)	0.30 (1,44)	–0.17 (–0,82)
4	*–0.16* (–0,88)	0.16 (0,82)	–0.21 (–1,02)
5	*–0.13* (–0,75)	–0.04 (–0,20)	–0.29 (–1,44)
6	*0.03* (0,19)	0.00 (0,02)	–0.19 (–0,98)
7	*0.27* (1,62)	0.04 (0,25)	–0.20 (–1,07)
8	*0.56 (3,34)*	0.13 (0,71)	–0.11 (–0,59)
9	*0.85 (4,62)*	0.30 (1,56)	–0.03 (–0,16)
High	*1.53 (6,78)*	*0.88 (3,62)*	–0.16 (–0,86)
# Sig	5	3	0
Adj R^2	0,69	0,74	0,78
χ^2-test	0,00	0,00	0,69

TABLE 18

Alphas' estimates and t-statistics: momentum-deciles

	1-factor	2-factor	alternative
Loser	*–1.04 (–3,38)*	–0.28 (–1,01)	–0.30 (–0,98)
2	–0.45 (–1,94)	0.07 (0,33)	0.05 (0,23)
3	–0.13 (–0,71)	0.27 (1,57)	0.24 (1,25)
4	–0.01 (–0,08)	0.25 (1,42)	0.23 (1,26)
5	0.07 (0,45)	0.30 (1,71)	0.29 (1,65)
6	–0.01 (–0,04)	0.16 (0,91)	0.16 (0,90)
7	–0.06 (–0,41)	–0.03 (–0,16)	–0.06 (–0,33)
8	0.02 (0,11)	–0.04 (–0,22)	–0.08 (–0,45)
9	0.07 (0,35)	–0.09 (–0,42)	–0.20 (–0,96)
Winner	0.42 (1,65)	0.19 (0,68)	–0.13 (–0,47)
# Sig	1	0	0
Adj R^2	0,68	0,74	0,72
χ^2-test	0,00	0,00	0,01

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

2. A second distress factor

The average monthly dollar returns for the distress-decile portfolios for the period 1980-2000 were already shown in Figure 2. Recall that we saw a monotone positive but S-shaped schedule where the highest-distress decile really pops out, which might indicate an extra distress risk. To resolve this we look at the alphas' *t*-statistics of ten distress decile portfolios for the one-factor CAPM and a two-factor version that includes the standard distress factor, the 30% top-B/M stocks minus the 30% bottom-B/M stocks (Table 17).

From Table 17 we conclude that the one-factor CAPM is not able to account for the distress risk: both growth and value portfolios have alphas that are significantly different from zero, with *t*'s ranging between –3 and +7. Table 17 also demonstrates that adding a standard distress factor portfolio improves the fit, but without whittling down the alphas to insignificant levels. On the basis of Figure 2 and the *t*-statistics of the one and two-factor model we propose a model with two distress factor portfolios: (*i*) extreme distress risk, *i.e.* the risk inherent to the highest B/M stocks that cannot be accounted for by beta risk nor normal distress risk; (*ii*) normal distress risk in the spirit of Fama and French (1993) but redefined to reduce overlap with extreme distress risk. Specifically, we introduce an *eHML* factor reflecting extreme risk, the return on a zero-investment portfolio that is long the highest B/M-decile stocks (*i.e.* hi-distress firms in decile ten) and short all other B/M deciles. We also work with *rHML* reflecting the regular distress risk, measured as the return on a zero-investment portfolio that is long the value stocks in B/M deciles 8 and 9 and short the growth stocks B/M deciles 1 and 2. Thus, the two-factor B/M model has become a three-factor model

$$R_i - r_j = a_i + b_i\,(R_m - r_f) + f_i eHML + j_i rHML + e_i \qquad (3)$$

Table 17 shows that our model does a good job in pricing the ten distress decile portfolios as none of the alphas are significantly different from zero. These conclusions remain valid when we use only NYSE stocks to calculate the decile breakpoints.

3. A modified momentum factor

The average monthly dollar returns of the momentum-deciles for the period 1980-2000 were already shown in Figure 3. We described the plot as an S-shaped rise. The apparent nonlinearity may still be picked up by the difference between decile membership and exposure, or by beta, so we investigate the ability of the one-factor CAPM to price unmanaged momentum portfolios.

From Table 18 we conclude, familiarly, that the one-factor CAPM does not fully capture momentum risk. We introduce a standard momentum factor portfolio as the difference between the 30% winners and 30% losers as in Rouwenhorst (1999), Carhart (1997) and Jegadeesh and Titman (1993). This model seems to work well: all individual alphas are insignificant. But in a more powerful test where we look at all risks simultaneously the standard momentum still badly misprices two portfolios (Table 19, to be discussed below). It turns out that no second momentum portfolio is needed to mend this. Rather, it suffices to redefine *WML* as the difference between returns from the 10% winners and the 20% losers.[13] Thus, our two-factor momentum model is:

$$R_i - r_f = a_i + b_i\,(R_m - r_f) + q_i WML + e_i \qquad (4)$$

Table 18 shows that our model does a good job in pricing the ten momentum decile portfolios as none of the alphas are significantly different from zero. These conclusions are also valid when we use only NYSE stocks to calculate the decile breakpoints.

In the next section we combine the three alternative models to one multi-factor model and test it more formally.

D. *Tests of the proposed factor specification*

In this section we combine the alternative size-, distress- and momentum model into one alternative multi-factor model. We show that, in pricing different kinds of unmanaged portfolios, this model does a better job than the standard multi-factor model with factor portfolios like in Fama and French ((1993), (1995), (1996a), (1996b)), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999). The standard multi-factor model is

$$\begin{aligned} R_i - r_f = {} & a_i + b_i\,(R_m - r_f) + g_i SMB + d_i HML + \\ & f_i WML + e_i \end{aligned} \qquad (5)$$

where *SMB* (small minus big) is the size factor portfolio, *viz.* a zero-investment portfolio that is long the 50% smallest stocks and short the 50% largest stocks; *HML* (high minus low) is the distress factor portfolio, a zero-investment portfolio that is long the 30% highest B/M stocks and short the 30% lowest B/M stocks; and *WML* (winner minus loser) is the momentum factor portfolio, a zero-investment portfolio that is long the 30% top past-performers (winners) and short the 30% lowest past-performers (losers). All portfolios are equally weighted and updated monthly.

Our alternative multi-factor model uses the same factor portfolios (size, distress and momentum) but is composed differently. Combining the alternative factor portfolios from the preceding sections into one model gives the following alternative multi-factor model

$$\begin{aligned} R_i - r_f = a_i + b_i\,(R_m - r_f) + g_i mSMB + d_i rSMB \\ + f_i eHML + j_i rHML + q_i WML + e_i \end{aligned} \tag{6}$$

with the factors as defined in Section III.C.

1. The pricing of one-dimensional test-portfolios

We demonstrate that the alternative model is a better model to price unmanaged one-dimensional test-portfolios – that is, stocks sorted on either size, distress or momentum – than the standard four-factor model.

In Table 19 the estimated alphas of the alternative model are always insignificantly different from zero, whereas under the standard four-factor model six alphas, in total, are clearly non-zero. We conclude that the alternative model is a better model in pricing unmanaged size-, distress- or momentum sorted portfolios.

2. The pricing of two-dimensional test-portfolios

We next demonstrate that the alternative model also is a better model to price the unmanaged two-dimensional size-distress-sorted test-portfolios that fared so badly in the FF tests of Section 1. We calculate the left-side portfolios in the same three ways that we used before: narrow-based breakpoints and test portfolios that use only stocks with price *and* book info; broad-based breakpoints and test portfolios where the maximum amount of return info is used; and the intermediate case, narrow-based breakpoints and broad-based test portfolios. The factor portfolios, in contrast, are always broad-based, again as described in Section II.

In Table 20, showing the results from the narrow-based break points and test portfolios, the alternative model has just half of the number of rejections that the standard version has (5 against 9). In Table 21

TABLE 19

Alpha estimates and-t-statistics: size, distress- and momentum sorted test-portfolios

	Size test-portfolios		Distress test-portfolios		Momentum test-portfolios	
	4-factor	Alternative	4-factor	Alternative	4-factor	Alternative
1	*1.32 (10.00)*	0.22 (0.84)	*0.44 (2.92)*	0.43 (1.69)	–0.24 (–1.10)	–0.05 (–0.12)
2	–0.29 (–1.93)	0.19 (0.66)	*0.44 (2.76)*	0.44 (1.60)	–0.07 (–0.40)	0.16 (0.53)
3	–0.32 (–1.88)	0.32 (1.13)	0.27 (1.53)	0.23 (0.77)	0.16 (1.06)	0.23 (0.87)
4	–0.29 (–1.49)	0.22 (0.64)	0.10 (0.55)	0.29 (0.97)	–0.04 (–0.27)	0.18 (0.67)
5	–0.11 (–0.54)	0.17 (0.50)	–0.10 (–0.53)	0.39 (1.35)	–0.05 (–0.29)	0.36 (1.39)
6	0.05 (0.26)	0.30 (0.92)	0.02 (0.11)	0.46 (1.69)	–0.17 (–1.03)	0.26 (1.09)
7	0.17 (0.85)	0.38 (1.22)	0.09 (0.61)	0.39 (1.58)	*–0.35 (–2.21)*	0.15 (0.65)
8	0.12 (0.70)	0.22 (0.84)	0.11 (0.66)	0.42 (1.61)	*–0.33 (–2.05)*	0.17 (0.69)
9	0.01 (0.04)	0.11 (0.50)	0.24 (1.56)	0.49 (1.85)	–0.20 (–1.26)	0.00 (0.02)
10	–0.03 (–0.40)	0.05 (0.38)	*0.80 (4.59)*	0.39 (1.57)	0.38 (1.64)	0.05 (0.16)
# Sig	1	0	3	0	2	0
Adj R^2	0.86	0.82	0.84	0.81	0.85	0.81
χ^2-test	0.00	0.88	0.00	0.30	0.00	0.37

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

(broad-based break points but narrow-based test portfolios) the evidence is not good in terms of the number of rejections (8 against 7), but the *t*-statistics are not nearly as large. The biggest is 2.68, down from 6.86, and the average significant *t*-statistic is 2.31, down from 3.40. In Table 22, finally, which is based on the broad data, the number of rejections falls from 7 (including a 6.23) to just one lone 2.12. We conclude that the alternative model does a better job in pricing the size-distress portfolios that failed the FF tests of Section II.

3. Out-of-sample tests

Fama and French (1995) point out that spurious common variation might be induced when the regressor portfolios *SMB* and *HML* are constructed from the same stocks as the regressand test-portfolios.[14] To avoid this, they provide a test where the stocks in the left-hand-side portfolios are different from those on the right-hand side. Specifically, they split the data into two equal groups. One group provides the dependent value-weighted size-B/M test portfolios for the time-series regressions. The other is used to form explanatory factor portfolio returns. In a second test, the roles of the two groups in the regressions are reversed.

We proceed similarly. At the end of each month *t*, stocks are ranked alphabetically; the odd-numbered are classified as A's, the even-numbered become B's. Within each half we proceed as before. Test set A provides the test portfolios for factor set B, and test set B similarly provides the test portfolios for factor set A.

From Table 23 and 24, we see that regressing A on B and *vice-versa* produces reassuringly similar alpha estimates and *t*-statistics, whether we compare the two tests of the 4-factor model or the two tests of the alternative model. Thus, spurious correlation does not seem to have been behind our earlier good results. This is in line with Fama and French (1995). The tables also show the importance of sample size: in the 4-factor model, the rejections are down from seven to five or four; still, the alternative model comes out with just one or two.

4. Industry test-portfolios

In this section we demonstrate that the alternative model does better than the standard four-factor model in pricing 36 unmanaged industry-sorted portfolios. The companies are classified according the Level 4 Datastream Industry Classification.

TABLE 20

Narrow-based breakpoints and test portfolios

		Low	2	3	4	High	
4-factor model	Small	*1.00 (3.84)*	0.28 (1.16)	0.28 (1.17)	0.23 (1.08)	*0.96 (5.84)*	
	2	–0.13 (–0.55)	0.43 (1.46)	0.10 (0.36)	*0.43 (2.13)*	0.30 (1.47)	
	3	*0.55 (2.10)*	0.50 (1.91)	0.16 (0.67)	0.01 (0.06)	*0.48 (2.05)*	# Sig: 9
	4	*0.45 (2.00)*	0.27 (1.32)	–0.21 (–1.12)	–0.02 (–0.12)	0.20 (1.15)	Adj R^2: 0.79
	Big	*0.46 (3.90)*	*–0.25 (–2.02)*	*–0.45 (–3.03)*	–0.14 (–1.01)	0.23 (0.93)	χ^2-test: 0.00 (143.75)
Alternative model	Small	0.81 (1.59)	–0.22 (–0.44)	*0.86 (2.29)*	0.67 (1.93)	*0.70 (2.49)*	
	2	0.03 (0.06)	0.58 (1.25)	*0.91 (2.25)*	*0.95 (3.17)*	0.44 (1.46)	
	3	0.39 (0.92)	0.49 (1.24)	0.65 (1.83)	0.36 (1.16)	0.59 (1.67)	# Sig: 5
	4	0.32 (0.99)	0.43 (1.39)	0.04 (0.14)	–0.08 (–0.29)	0.33 (1.18)	Adj R^2: 0.76
	Big	*0.51 (2.67)*	–0.08 (–0.41)	–0.27 (–1.18)	0.02 (0.10)	–0.18 (–0.65)	χ^2-test: 0.00 (67.60)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 21

Broad-based breakpoints; narrow-based test-portfolios

		Low	2	3	4	High	
4-factor model	Small	*2.01 (3.22)*	0.60 (1.39)	0.33 (0.92)	0.40 (1.60)	*1.17 (6.86)*	
	2	0.33 (1.32)	–0.01 (–0.05)	–0.04 (–0.15)	0.10 (0.48)	0.36 (1.79)	
	3	0.05 (0.19)	*0.56 (2.29)*	0.14 (0.49)	0.21 (1.00)	*0.58 (2.49)*	# Sig: 7
	4	*0.62 (2.66)*	0.26 (1.15)	0.03 (0.13)	0.09 (0.48)	0.32 (1.68)	Adj R^2: 0.82
	Big	*0.39 (3.53)*	–0.16 (–1.24)	*–0.40 (–2.75)*	–0.13 (–1.00)	0.15 (0.77)	χ^2-test: 0.00 (173.12)
Alternative model	Small	*2.04 (2.36)*	–0.94 (–1.32)	0.38 (0.74)	0.72 (1.69)	*0.76 (2.56)*	
	2	0.43 (0.89)	–0.01 (–0.01)	*0.96 (2.20)*	*0.89 (2.68)*	0.48 (1.41)	
	3	–0.13 (–0.33)	*0.79 (1.99)*	*0.81 (2.06)*	*0.68 (2.18)*	0.54 (1.55)	# Sig: 1
	4	0.51 (1.35)	0.38 (1.12)	0.39 (1.32)	0.16 (0.55)	0.43 (1.64)	Adj R^2: 0.79
	Big	*0.42 (2.45)*	0.02 (0.11)	–0.17 (–0.79)	–0.13 (0.64)	–0.03 (–0.15)	χ^2-test: 0.0007 (53.70)

TABLE 22

Broad-based breakpoints and test-portfolios

		Low	2	3	4	High	
4-factor model	Small	*0.70 (3.11)*	*0.58 (2.48)*	*0.57 (2.97)*	0.08 (0.43)	*0.80 (6.23)*	
	2	–0.25 (–1.16)	–0.15 (–0.58)	–0.32 (–1.53)	*–0.45 (–2.22)*	–0.22 (–1.32)	
	3	0.07 (0.26)	0.20 (0.89)	–0.25 (–1.01)	–0.22 (–1.08)	0.13 (0.61)	# Sig: 7
	4	0.47 (1.90)	0.08 (0.38)	0.02 (0.08)	–0.02 (–0.09)	0.01 (0.06)	Adj R^2: 0.82
	Big	*0.37 (3.37)*	–0.17 (–1.30)	*–0.33 (–2.34)*	–0.20 (–1.69)	0.16 (0.84)	χ^2-test: 0.00 (173.12)
Alternative model	Small	–0.07 (–0.18)	–0.14 (–0.38)	0.30 (0.97)	–0.08 (–0.26)	0.55 (1.96)	
	2	0.09 (0.25)	0.04 (0.09)	0.52 (1.54)	0.33 (1.08)	0.27 (0.90)	
	3	–0.14 (–0.36)	0.40 (1.08)	0.39 (1.09)	0.33 (1.07)	0.22 (0.61)	# Sig: 1
	4	0.28 (0.76)	0.24 (0.74)	0.39 (1.32)	0.29 (1.08)	0.37 (1.33)	Adj R^2: 0.79
	Big	*0.36 (2.12)*	0.00 (–0.00)	0.00 (0.01)	–0.14 (–0.66)	0.11 (0.52)	χ^2-test: 0.0007 (53.70)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 23

Separate data left and right; broad-based breakpoints and test-portfolios, A on B

		Low	2	3	4	High	
4-factor model	Small	0.54 (1.94)	0.38 (1.38)	0.18 (0.65)	*0.67 (2.59)*	*0.36 (2.49)*	
	2	–0.28 (–0.97)	–0.16 (–0.51)	–0.16 (–0.69)	–0.07 (–0.30)	*–0.46 (–3.18)*	
	3	–0.36 (–1.28)	–0.26 (–0.95)	–0.31 (–1.25)	–0.17 (–0.75)	*–0.44 (–2.71)*	# Sig: 5
	4	0.03 (0.11)	–0.24 (–1.11)	–0.17 (–0.80)	0.17 (0.83)	–0.08 (–0.50)	Adj R^2: 0.76
	Big	*0.92 (4.28)*	0.12 (0.52)	0.40 (1.92)	0.20 (0.96)	0.02 (0.12)	χ^2-test: 0.00 (130.76)
Alternative model	Small	0.18 (0.42)	–0.09 (–0.22)	–0.30 (–0.70)	0.28 (0.68)	0.18 (0.74)	
	2	–0.54 (–1.24)	–0.25 (–0.53)	0.04 (0.12)	–0.12 (–0.33)	–0.25 (–1.00)	
	3	–0.19 (–0.47)	0.15 (0.38)	0.06 (0.18)	0.23 (0.74)	–0.33 (–1.40)	# Sig: 1
	4	–0.06 (–0.18)	0.11 (0.36)	0.09 (0.30)	0.43 (1.62)	0.07 (0.33)	Adj R^2: 0.71
	Big	0.56 (1.76)	0.16 (0.45)	0.34 (1.05)	*0.54 (2.05)*	–0.21 (–0.78)	χ^2-test: 0.0003 (57.02)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 24

Separate data left and right; broad-based breakpoints and test-portfolios, B on A

		Low	2	3	4	High	
4-factor model	Small	0.43 (1.46)	0.67 (1.92)	0.26 (1.09)	*0.50 (2.02)*	*0.56 (3.67)*	
	2	–0.34 (–1.20)	–0.11 (–0.39)	0.25 (0.95)	–0.14 (–0.59)	–0.24 (–1.46)	
	3	0.07 (0.26)	–0.27 (–1.09)	–0.11 (–0.47)	–0.15 (–0.68)	*–0.55 (–3.03)*	# Sig: 4
	4	0.02 (0.08)	–0.13 (–0.57)	0.19 (0.91)	–0.07 (–0.37)	–0.16 (–1.08)	Adj R^2: 0.76
	Big	*0.98 (5.73)*	0.25 (1.37)	0.26 (1.16)	0.40 (2.21)	0.00 (0.02)	χ^2-test: 0.00 (159.71)
Alternative model	Small	0.17 (0.44)	–0.10 (–0.25)	–0.45 (–1.20)	0.10 (0.25)	0.32 (1.49)	
	2	–0.50 (–1.24)	–0.41 (–1.05)	0.17 (0.49)	–0.14 (–0.42)	–0.17 (–0.84)	
	3	0.25 (0.73)	0.08 (0.24)	0.06 (0.16)	0.06 (0.21)	–0.33 (–1.57)	# Sig: 2
	4	0.39 (1.19)	0.20 (0.61)	0.45 (1.51)	0.01 (0.03)	0.03 (0.15)	Adj R^2: 0.73
	Big	*1.00 (3.56)*	0.36 (1.37)	0.31 (1.10)	*0.65 (2.60)*	0.04 (0.16)	χ^2-test: 0.00 (85.25)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

From Table 25, we see that the alternative model produces fewer significant alphas than the standard four-factor model, five as opposed to thirteen. We conclude that the alternative model does a better job in pricing unmanaged industry portfolios. The remaining significant unexplained returns in the alternative may reflect the unusual performance of certain industries at the end of the nineties (e.g. ICT- and biotech bubbles). It looks likely that these significant alphas would disappear if one extends the test period beyond 2000.

E. *Conclusion*

The alternative model provides a significantly improved version relative to the standard three- or four-factor CAPM model with factor portfolios of Fama and French ((1993), (1995), (1996a), (1996b)), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999). Our model produces estimated alphas closer to zero for one-dimensional size-, distress-, momentum- and industry portfolios and two-dimensional size-distress portfolios. However, the evidence so far bears on the US-market only, and the factor portfolios were hand-picked to fit this data set. In the next section, we accordingly test whether the alternative factor portfolios keep on producing estimated alphas close to zero in an international setting.

IV. INTERNATIONAL VALIDATION

Recall that Fama and French (1993) calculated the size- and distress-decile breakpoints on the NYSE stocks only, and used these to catalogue all US stocks, including Amex and NASDAQ stocks. One of our criticisms was that this procedure is difficult to implement in an international setting. When Fama and French (1998) investigate value versus growth effects in an international setting they abandon this procedure and calculate the decile breakpoints from all stocks.

They proceed by calculating their size and distress factor portfolios (*SMB* and *HML*) for each country separately. The global *SMB* and *HML* factor portfolios are then constructed as averages of these country factor portfolios, weighted on the basis of the MSCI country weights. This surely avoids the risk that, say, the Big portfolio becomes very much a US affair. However, in some countries the range of corporate size or B/M is quite narrow: many small Western

TABLE 25

4-factor model vs. alternative model: alpha-t-statistics for industry portfolios

	4-factor model	Alternative model		4-factor model	Alternative model
aerosp. & def.	–0.21 (–0.83)	–0.03 (–0.08)	leisure & hotels	–0.45 (–1.78)	–0.17 (–0.45)
autom. & parts	–0.27 (–1.10)	–0.11 (–0.29)	life assurance	*–0.63 (–2.46)*	–0.25 (–0.68)
banks	*–0.80 (–3.85)*	–0.34 (–1.13)	media & entert.	*0.49 (2.22)*	0.38 (1.19)
beverages	–0.37 (–1.34)	–0.08 (–0.21)	mining	–0.28 (–0.46)	0.38 (0.44)
chemicals	–0.06 (–0.25)	0.22 (0.68)	oil & gas	–0.73 (–1.45)	0.21 (0.30)
constr. mats.	*–0.67 (–2.67)*	–0.26 (–0.68)	prsnl care & hse	–0.05 (–0.20)	0.57 (1.75)
divers. industry	0.21 (0.90)	0.62 (1.72)	pharmc & biotch	*1.27 (3.76)*	*1.26 (2.52)*
electricity	–0.40 (–1.65)	–0.19 (–0.59)	real estate	*–1.09 (–5.00)*	*–0.71 (–2.22)*
electro & electic	*0.62 (2.41)*	0.52 (1.26)	retailer (general)	0.00 (0.01)	0.17 (0.35)
engin. & machin.	–0.17 (–0.82)	0.01 (0.04)	softwr & services	*1.51 (5.04)*	*1.16 (2.56)*
food & drug ret.	–0.32 (–1.24)	0.37 (1.02)	specialty & finan	*–0.74 (–2.60)*	–0.61 (–1.49)
food producers	–0.06 (–0.31)	*0.60 (2.39)*	steel& oth.metal	–0.54 (–1.71)	–0.50 (–1.08)
forestry & paper	–0.30 (–0.94)	–0.37 (–0.80)	support services	0.35 (1.61)	*0.74 (2.11)*
hshld gd & textil	–0.19 (–0.78)	0.32 (0.86)	telecom services	*0.94 (3.16)*	0.69 (1.61)
healthcare	*0.54 (1.99)*	0.81 (1.96)	tobacco	0.95 (1.40)	1.26 (1.34)
i/t hardware	*1.45 (4.11)*	0.81 (1.52)	transport	–0.35 (–1.44)	0.08 (0.23)
insurance	*–0.67 (–2.90)*	–0.21 (–0.66)	other utilities	–0.24 (–1.21)	0.27 (1.00)
			# Sig	13	5
			Adj R	64.89	61.94
			χ^2-test	0.00 (176.83)	0.01 (58.69)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

countries have no really big firms, and some emerging markets specialize in one sector, thus reflecting the rather similar size or book-to-market figures that are typical for that industry. In short, one issue is whether classification into, say, the B or S buckets should be done country by country or via one global list.

Another issue again is value weighting. In FF (1998) this happens within countries and across countries, via the MSCI weights. The combined effect is to downplay the small-firm effect even more than within the US study: many smallish firms are conjured away, being classified as locally Big rather than globally Small and then drowned in the value-weighted global Big portfolio. A related drawback is that both S and B are now dominated by US firms, making the international sample rather similar to the American one. Lastly, in FF (1998) there is a requirement that book data be known, and this again eliminates many of the smaller stocks. We accordingly prefer to construct the size-, distress- and momentum factor portfolios in one shot, from the global stock list; we use equally weighted portfolio returns for all factors other than the world market; and whenever possible we include also stocks with an unknown market value or book-to-market value.

A description of the international database can be found in the appendix--available on request. For current purposes it suffices to note that the international database covers 39 countries,[15] both developed and emerging. In building the country list we tried to cover as much of the world as possible taking into account the availability and reliability of data. For each stock we know the end-of-month monthly dollar return, monthly dollar market value, level-4 industry category, nationality, and (sometimes) monthly book-to-market ratio, for part or all of the period 1980-2000. For each country we also know the end-of-month monthly dollar exchange rate return and end-of-month monthly risk-free rates for the period 1980-2000. Most of these exchange and interest rates originate from the IFS-IMF database, and in general the end-of-period central bank discount rate is taken as the risk-free interest rate. Like in previous sections, all portfolios are equally weighted and updated monthly. The world market return is the monthly dollar return of DataStream's world-market index.

In the next paragraphs, we investigate whether the alternative composition of the factor portfolios from the US setting (previous section), also works in an international setting. To do so, we compare the alphas' estimates and *t*-statistics of five models for 103 test portfolios, constructed using five alternative criteria: (*i*) ten size deciles;

(*ii*) ten distress deciles; (*iii*) ten momentum deciles; (*iv*) 39 industry portfolios; (*v*) and 34 country portfolios.

The first model is the basic one-factor CAPM,

$$R_{t+1} - r_{f,t}^{US} = a + b\left(R_{m,t+1} - r_{f,t}^{US}\right) + e_{t+1} \tag{7}$$

When testing the pricing of one-dimensional style portfolios, one obvious generalisation is to add either the standard size-, distress- or momentum factor. That is,

$$R_{t+1} - r_{f,t}^{US} = a + b\left(R_{m,t+1} - r_{f,t}^{US}\right) + gRF_{t+1} + e_{t+1} \tag{8}$$

where RF_{t+1} is the appropriate factor portfolio corresponding to the criterion used when sorting the left-hand-side test portfolios. The factor portfolios are set up following Fama and French ((1993), (1995), (1996a), (1996b)), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999) except that portfolios are equally weighted and updated monthly.[16]

CAPM 3 is the nested version of the above three, the standard four-factor CAPM:

$$\begin{aligned} &R_{t+1} - r_{f,t}^{US} = \\ &a + b\left(R_{m,t+1} - r_{f,t}^{US}\right) + gSMB_{t+1} + dHML_{t+1} + fWML_{t+1} + e_{t+1} \end{aligned} \tag{9}$$

The next candidate model adds the InCAPM exchange-rate factors to this four-factor specification.[17] Individual stocks can be exposed to exchange rate fluctuations because the going exchange rate partly determines a firm’s domestic-currency cash flows from importing or exporting goods or services. Exchange rates also affect a firm’s foreign-currency prices for goods or services and hence the demand for its output. Lastly, the firm’s stock price is converted into dollars; thus, even if a firm’s own-currency stock return would not be exposed to exchange rate fluctuations, its translated stock return would still be. This last aspect is the main reason why an InCAPM may be needed when testing the pricing of stocks from many countries. US stocks are also held by non-US residents, so in principle the average investor does care about exposures. But for US stocks it is hard to establish that currency exposures are in fact non-zero (see Bartov and Bodnar (1994); Bodnar and Gentry (1993); Allayannis (1995); and Allayannis (1997)). In contrast, foreign stocks on average do betray their nationality via exposure to their own exchange rate (Adler and Dumas (1984)).

The Solnik-Sercu model is a static CAPM without state variables – so an obvious extension will be to add the standard *SMB*, *HML* and Momentum factors – featuring the world market-portfolio return and the excess returns from investing in each non-USD currency. Including all 39 currencies is not recommendable as the power of the alpha tests will drop dramatically, but apart from this consideration there are no clear guidelines or standard practices. Jorion (1990) proposes to use a fixed trade-weighted basket of currencies, but this assumes that all stocks have a vector of currency exposures that is proportional to the trade weights – a restriction which Rees and Unni (1999) reject empirically. We adopt a compromise. Specifically, we include in every regression the individual currencies of seven countries (C7), taking at least one currency per continent and looking, per continent, at ecoonomic weight and number of stocks in our data base. This "C7" list contains the Canadian Dollar, British Pound and Deautsche Mark, Japanese Yen and Korean Won, Australian Dollar and South African Rand. All stocks are allowed to be exposed, without any prior restrictions, to each of these C7 currencies. On top of that, non-C7 stocks are assumed to have a common exposure to their own exchange rate (Adler and Dumas (1984)). Since the regressand variables are portfolio returns, this last assumption means that for each such test portfolio a basket of currency deposits is created which gives to each non-C7 currency the same weight as the stocks from that country have in the particular test portfolio. Thus, if a portfolio contains n_1 stocks from non-C7 currency 1 and n_2 stocks from non-C7 currency 2, then the basket consists of $n_1/\ (n_1+n_2)$ invested in currency 1 and $n_2/\ (n_1+n_2)$ invested in currency 2. In regressions with country portfolios as the regressands, this currency factor collapses to the country's own currency factor; for the C7 country indices and U.S. index, this 8th currency factor is therefore redundant and dropped from the regression. The nested Solnik-Sercu/4-factor model reads like

$$\begin{aligned} R_{t+1} - r_{f,t}^{US} = a_i + b_i\left(R_{m,t+1} - r_{f,t}^{US}\right) + g_i SMB_{t+1} + d_i HML_{t+1} + \\ f_i WML_{t+1} + \sum_{k=1}^{7} y_{i,k} XF_{k,t+1} + z_i CXF_{i,t+1} + e_{t+1} \end{aligned} \tag{10}$$

$$\text{and} \quad XF_{t+1} = s_{t+1} + r_{f,t}^{*} - r_{f,t}^{US} \quad \text{with} \quad s_{t+1} = \frac{S_{t+1} - S_t}{S_t} \tag{11}$$

where the right-side factor portfolios (SMB, HML and WML) are like in the standard four-factor model. Subscript i stands for the i-th left-side test-portfolio, subscript k denotes the k-th exchange factor portfolio and CXF_i refers to the compound non-C7 exchange factor portfolio tailored for the i-th test-portfolio. S denotes the going spot exchange rate (USD per foreign currency) and $r^{*}_{f,t}$ the foreign risk-free interest rate.

The last CAPM candidate is obtained by adding the two factors identified in the US tests (the micro-stock and extreme-distress factors) and re-specifying the SMB and HML factors as described in Section III:

$$\begin{aligned} R_{t+1} - r^{US}_{f,t} = \alpha_i + \beta_i\left(R_{m,t+1} - r^{US}_{f,t}\right) + \gamma_i mSMB_{t+1} + \delta_i rSMB_{t+1} + \phi_i eHML_{t+1} + \\ \varphi_i rHML_{t+1} + \theta_i WML_{t+1} + \sum_{k=1}^{7} \psi_{i,k} XF_{k,t+1} + \zeta_i CXF_{i,t+1} + \varepsilon_{t+1} \end{aligned} \tag{12}$$

A. *Results for size, distress, and momentum test-portfolios*

US firms take up 55 percent of the total sample by numbers, but are relatively underrepresented in the lower size quintile, where they provide only 45 percent of the observations. Still, when we look at the plot, in Figure 16, of the average monthly dollar return of the ten international size-deciles for the period 1980-2000, it looks a lot like Figure 1 (US market). Thus, also in an international setting there seems to exist a strong small firm effect for the smallest stocks – unless, of course, the high average return would be explained by beta. In contrast, the inverted small-firm effect in the US market, where the average return of the biggest firms was slightly higher than the average-sized firms, disappears in an international setting: bigger firms earn monotonely less. From Table 26, we see that also in an international setting the specification of the factor portfolios plays an important role in the ability of an asset pricing model to price unmanaged portfolios. The alpha-t-statistics in Table 26 show that the specification adopted in the US study of the preceding section also produces insignificant alphas for the ten size deciles in an international setting.

We now turn to Figure 17 which plots the average monthly dollar return for ten international B/M decile portfolios for the period 1980-2000. Again Figure 17 resembles its US counterpart, Figure 2, which exhibits a gradually rising monthly average return as we move from

FIGURE 16

Average monthly dollar return for ten international distress-sorted deciles

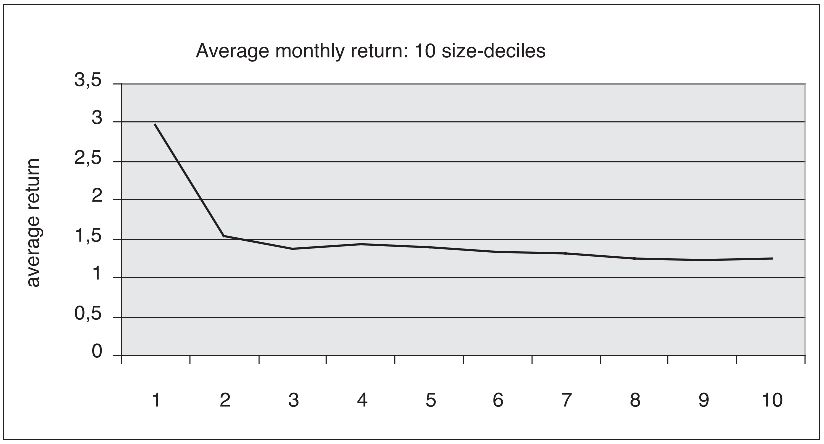
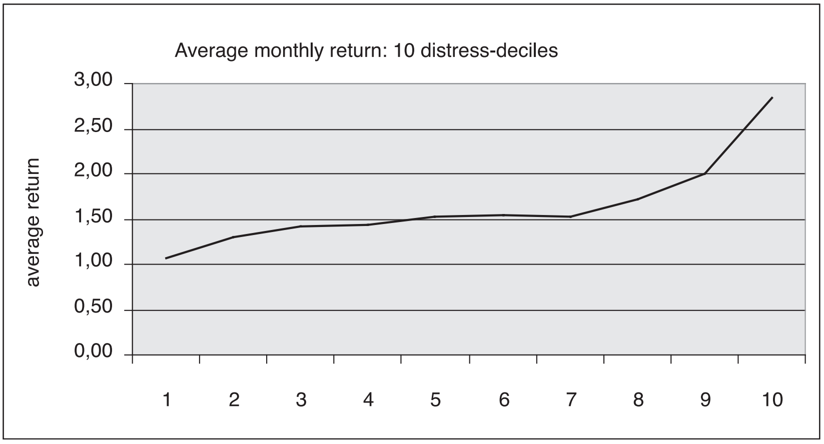


FIGURE 17

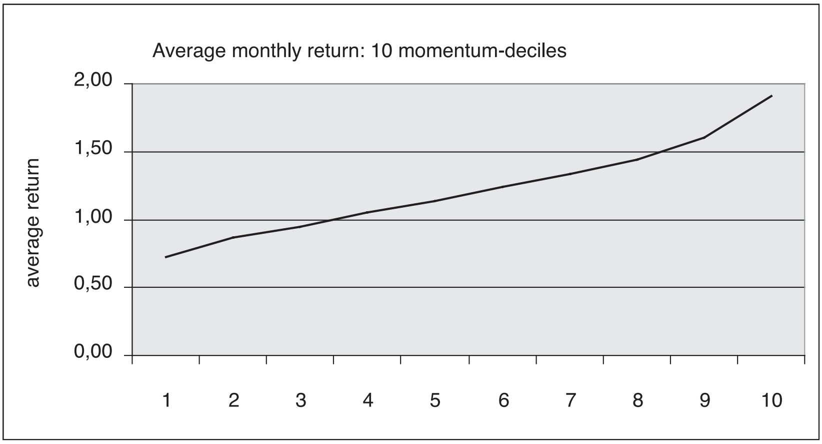
Average monthly dollar return for ten international momentum-sorted deciles



growth stocks (low B/M value) to distress or value stocks (high B/M value). There is an S-shape, and especially the highest distress decile pops out again. From Table 27, we conclude that also in an international setting, the generalized model seems to outperform the standard international model in pricing unmanaged B/M based test-portfolios. Only the third B/M decile portfolio remains significant.

FIGURE 18

Average monthly dollar return for ten international momentum-sorted deciles



Lastly, Figure 18 plots the average monthly dollar return for ten international momentum decile portfolios for the period 1980-2000. Remember that Figure 3 (US market) looked like an S-shaped positive schedule not too far from linearity. Figure 18 resembles a linear rise even more. Apparently, in an international setting, the average return of international momentum portfolios rises at a more constant rate. From Table 28, we see that the standard two-factor momentum capm, produces the best results, delivering zero rejections against one marginal t=2.01 for the modified model. However, the difference is small, and especially the standard 4-factor model does poorly. More fundamentally, the standard two-factor momentum CAPM does best only once (and marginally at that), notably when pricing momentum-based test-portfolios; in all other applications it flaunders badly, while the proposed multi-factor CAPM remains superior in pricing unmanaged momentum based international test-portfolios.

B. *Results for industry and country test portfolios*

In this paragraph, we first try to price unmanaged international industry portfolios. We follow the Leve-4 Datastream Industry Classification which contains 34 meaningful industries.

TABLE 26

Alphas' estimates and t-statistics: size-portfolios

	Standard one-factor	Standard two-factor	Standard four-factor	Standard InCapm	Alternative InCapm
Small	*1.87 (8.59)*	*0.80 (8.51)*	*0.96 (9.22)*	*0.95 (9.03)*	0.12 (0.63)
2	0.37 (1.79)	*–0.63 (–6.69)*	*–0.51 (–4.84)*	*–0.50 (–4.71)*	–0.01 (–0.03)
3	0.20 (1.07)	*–0.64 (–6.15)*	*–0.54 (–4.63)*	*–0.55 (–4.73)*	0.18 (0.86)
4	0.21 (1.19)	*–0.47 (–3.75)*	*–0.34 (–2.49)*	*–0.38 (–2.78)*	0.23 (1.00)
5	0.15 (0.98)	*–0.33 (–2.48)*	–0.19 (–1.29)	–0.22 (–1.51)	0.26 (1.12)
6	0.06 (0.42)	*–0.30 (–2.21)*	–0.13 (–0.86)	–0.16 (–1.08)	0.12 (0.50)
7	0.04 (0.33)	–0.18 (–1.34)	0.04 (0.28)	0.02 (0.13)	0.17 (0.79)
8	–0.06 (–0.50)	–0.20 (–1.61)	0.03 (0.19)	0.02 (0.16)	0.07 (0.34)
9	–0.12 (–1.17)	–0.18 (–1.72)	–0.05 (–0.45)	–0.06 (–0.48)	–0.02 (–0.11)
Big	–0.08 (–1.49)	–0.10 (–1.67)	–0.06 (–0.91)	–0.06 (–0.93)	0.06 (0.67)
# Sig	1	6	4	4	0
Adj R^2	0.70	0.85	0.85	0.85	0.81
p-value, χ^2-test	0.00 (404)	0.00 (337)	0.00 (283)	0.00 (273)	0.12 (15.36)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 27

Alphas' estimates and t-statistics: distress-portfolios

	Standard one-factor	Standard two-factor	Standard four-factor	Standard InCapm	Alternative InCapm
Small	–0.30 (–1.73)	0.16 (0.99)	–0.09 (–0.64)	–0.05 (–0.35)	0.09 (0.48)
2	–0.09 (–0.56)	*0.32 (2.20)*	0.23 (1.78)	0.24 (1.86)	0.37 (1.83)
3	0.06 (0.44)	*0.35 (2.57)*	*0.31 (2.32)*	*0.30 (2.22)*	*0.43 (2.04)*
4	0.10 (0.74)	0.25 (1.79)	0.24 (1.67)	0.21 (1.47)	0.31 (1.41)
5	0.22 (1.59)	0.23 (1.58)	0.19 (1.23)	0.15 (0.97)	0.35 (1.54)
6	*0.27 (1.99)*	0.14 (1.00)	0.10 (0.69)	0.07 (0.5)	0.27 (1.25)
7	*0.30 (2.15)*	0.01 (0.03)	–0.09 (–0.67)	–0.14 (–0.97)	0.07 (0.35)
8	*0.54 (3.54)*	0.08 (0.61)	–0.01 (–0.11)	–0.04 (–0.28)	0.25 (1.29)
9	*0.83 (4.61)*	0.18 (1.31)	0.03 (0.21)	–0.01 (–0.06)	0.19 (0.95)
Big	*1.60 (5.85)*	*0.57 (2.87)*	*0.44 (2.55)*	*0.54 (3.25)*	0.26 (1.31)
# Sig	5	3	2	2	1
Adj R^2	0.69	0.78	0.83	0.84	0.82
p-value, χ^2-test	0.00 (56.87)	0.001 (29.71)	0.00 (43.59)	0.00 (47.30)	0.05 (18.63)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 28

Alphas' estimates and t-statistics: momentum portfolios

	Standard one-factor	Standard two-factor	Standard four-factor	Standard InCapm	Alternative InCapm
Small	*–0.59 (–2.28)*	0.17 (0.87)	*–0.59 (–3.85)*	*–0.57 (–3.67)*	*–0.53 (–2.01)*
2	*–0.39 (–2.02)*	0.21 (1.55)	–0.18 (–1.42)	–0.19 (–1.43)	0.04 (0.17)
3	–0.27 (–1.74)	0.14 (1.09)	–0.23 (–1.91)	–0.24 (–1.93)	0.05 (0.23)
4	–0.14 (–1.00)	0.12 (0.95)	*–0.28 (–2.25)*	*–0.29 (–2.32)*	0.03 (0.14)
5	–0.05 (–0.38)	0.10 (0.77)	*–0.30 (–2.39)*	*–0.32 (–2.56)*	–0.03 (–0.14)
6	0.06 (0.49)	0.10 (0.78)	*–0.31 (–2.52)*	*–0.33 (–2.70)*	–0.02 (–0.11)
7	0.16 (1.28)	0.11 (0.89)	*–0.28 (–2.28)*	*–0.30 (–2.46)*	0.04 (0.19)
8	0.23 (1.83)	0.10 (0.81)	*–0.32 (–2.70)*	*–0.33 (–2.79)*	0.00 (0.02)
9	*0.37 (2.63)*	0.15 (1.10)	*–0.31 (–2.53)*	*–0.31 (–2.53)*	–0.03 (–0.18)
Big	*0.58 (3.00)*	0.27 (1.46)	*–0.37 (–2.48)*	*–0.34 (–2.29)*	–0.24 (–1.05)
# Sig	4	0	8	8	1
Adj R^2	0.70	0.77	0.85	0.85	0.81
p-value, χ^2-test	0.002 (27.79)	0.55 (8.79)	0.00 (37.00)	0.00 (32.31)	0.004 (25.99)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

From Table 29 we conclude that the alternative international CAPM, with the two new factors, is the best model to price unmanaged international industry portfolios. Compared with Table 25 (the industry-portfolio test for US market) we notice that the ICT- and Biotech-bubble of the late 90's appears to be mostly a US issue. In an international setting 1980-2000, the abnormal positive returns noted for many US sectors – notably food producers, pharmaceuticals and biotechnology, support services, software and services industries – are not present or at least not significant. The negative abnormal return of the real-estate industry remains present in the international setting, though.

In this paragraph, we also try to price 39 unmanaged country portfolios. From Table 30, it is difficult to tell whether for the purpose of pricing unmanaged country portfolios the four-factor model is best or the extended InCAPM: the latter is marginally beaten in terms of the number of rejections, but does better in the sense that the t-ratios remain smaller. Interestingly, according to the standard four-factor and standard international model, the US market's performance was not impressive – to the contrary, in fact – while Japan's was; our proposed generalisation of international model is agnostic about both.

We conclude that also in an international setting the specific composition of the factor portfolios plays an important role in the ability of an asset pricing model to price unmanaged portfolios. We showed that the same composition that performed well in the US-market also works in an international setting.

V. CONCLUSION

We showed that some aspects of a CAPM test can not be taken for granted as they have a significant impact on the test results. Elements that have an influence are: (*i*) the coverage of small stocks in the database; (*ii*) the frequency of portfolio updating (monthly or yearly); (*iii*) the way of weighting the portfolio returns (equally or value); (*iv*) the way of calculating factor portfolio-portfolios (broad- *v.* narrow-based risk-portfolios i.e. are stocks with unknown market value or book-to-market value included or not); (*v*) the way of calculating test portfolios (intersections of two classifications or unions; and, in the first case, what to do with the stocks without full information); (*vi*) the coverage of the list from which the size- and distress-decile

TABLE 29

Alphas' estimates and t-statistics for 34 international industry-portfolios

	Standard one-factor	Standard four-factor	Standard InCapm	Alternative InCapm
aerosp. & def.	0.12 (0.55)	*–0.62 (–3.02)*	*–0.67 (–3.38)*	–0.10 (–0.34)
autom. & parts	0.17 (0.96)	–0.09 (–0.49)	–0.17 (–1.04)	–0.25 (–1.01)
banks	*0.37 (2.68)*	–0.07 (–0.52)	*–0.26 (–2.04)*	–0.29 (–1.61)
beverages	*0.36 (2.59)*	0.17 (1.03)	–0.10 (–0.89)	–0.17 (–1.09)
chemicals	0.19 (1.10)	0.17 (0.90)	0.00 (–0.02)	–0.14 (–0.61)
constr. mats.	0.14 (0.80)	0.04 (0.21)	–0.13 (–0.77)	–0.33 (–1.36)
divers. industry	0.22 (1.34)	*–0.31 (–2.06)*	*–0.44 (–4.13)*	*–0.37 (–2.17)*
electricity	*0.45 (2.88)*	*0.43 (2.51)*	0.27 (1.84)	0.21 (1.01)
electro & electic	0.14 (0.69)	–0.16 (–0.84)	–0.20 (–1.13)	–0.19 (–0.65)
engin. & machin.	–0.03 (–0.21)	–0.21 (–1.24)	–0.28 (–1.81)	–0.43 (–1.79)
food & drug ret.	*0.33 (2.29)*	0.18 (1.11)	–0.12 (–0.84)	0.19 (0.89)
food producers	*0.40 (2.92)*	0.26 (1.76)	–0.04 (–0.34)	0.04 (0.21)
forestry & paper	0.07 (0.37)	–0.31 (–1.55)	*–0.39 (–2.44)*	–0.29 (–1.24)
hshld gd & textil	0.32 (1.88)	–0.03 (–0.15)	–0.19 (–1.27)	–0.13 (–0.55)
healthcare	0.43 (1.48)	–0.40 (–1.47)	–0.41 (–1.56)	0.23 (0.60)
i/t hardware	*0.62 (2.03)*	0.01 (0.04)	–0.10 (–0.36)	–0.28 (–0.68)
insurance	*0.30 (2.01)*	–0.09 (–0.52)	–0.26 (–1.94)	–0.31 (–1.65)
leisure & hotels	0.16 (0.91)	*–0.45 (–2.71)*	*–0.60 (–3.90)*	–0.42 (–1.75)
life assurance	*0.55 (2.97)*	0.03 (0.16)	–0.24 (–1.35)	–0.01 (–0.03)
media & entert.	*0.72 (4.25)*	0.16 (1.06)	–0.02 (–0.17)	0.00 (0.01)
mining	0.05 (0.12)	–0.35 (–0.75)	–0.57 (–1.30)	–0.46 (–0.71)

	Standard one-factor	Standard four-factor	Standard InCapm	Alternative InCapm
oil and gas	–0.24 (–0.73)	*–0.86 (–2.39)*	*–1.03 (–3.42)*	–0.33 (–0.75)
persnl care	*0.46 (2.68)*	0.01 (0.06)	–0.22 (–1.39)	–0.15 (–0.64)
pharmc & biotch	*0.68 (3.02)*	0.26 (1.20)	0.07 (0.31)	0.00 (–0.01)
real estate	0.04 (0.24)	–0.27 (–1.52)	*–0.59 (–4.95)*	*–0.64 (–3.53)*
retailer (general)	0.33 (1.88)	–0.11 (–0.63)	–0.24 (–1.39)	–0.30 (–1.11)
softwr & services	0.57 (1.68)	–0.06 (–0.20)	–0.11 (–0.44)	–0.08 (–0.18)
Specialty & finan	0.40 (1.94)	–0.34 (–1.67)	–0.37 (–1.82)	–0.58 (–1.93)
steel & metal	0.01 (0.04)	–0.32 (–1.43)	–0.30 (–1.59)	–0.49 (–1.74)
support services	0.32 (1.57)	–0.33 (–1.89)	*–0.52 (–3.20)*	–0.17 (–0.66)
telecom services	*0.77 (3.14)*	0.13 (0.54)	0.11 (0.53)	–0.02 (–0.06)
Tobacco	*1.00 (3.40)*	0.31 (0.93)	0.19 (0.84)	0.09 (0.28)
Transport	0.12 (0.83)	–0.01 (–0.06)	–0.15 (–1.10)	–0.12 (–0.59)
other utilities	*0.38 (2.80)*	0.19 (1.23)	0.08 (0.55)	0.23 (1.15)
# Sig	14	5	8	2
Adj R^2	0.57	0.65	0.74	0.72
x^2-test	0.00 (101)	0.00 (116)	0.00 (144)	0.002 (63)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

TABLE 30

Alphas estimates and t-statistics for 39 country portfolios

	Standard one-factor	Standard four-factor	Standard InCapm	Alternative InCapm
argentine	2.64 (1.54)	1.86 (0.92)	1.34 (0.67)	1.24 (0.46)
australia	0.47 (1.28)	–0.25 (–0.60)	–0.37 (–1.20)	–0.13 (–0.26)
germany	0.05 (0.19)	0.01 (0.03)	0.11 (0.39)	0.17 (0.47)
belgium	0.29 (1.05)	0.24 (0.73)	0.21 (0.68)	0.44 (1.06)
brazil	1.74 (1.06)	–0.27 (–0.14)	–0.51 (–0.22)	–0.46 (–0.15)
colombia	0.21 (0.23)	–0.76 (–0.68)	*–2.79 (–2.22)*	*–3.37 (–2.04)*
china	2.90 (1.84)	2.47 (1.25)	1.46 (0.72)	3.13 (1.21)
chili	*1.86 (2.78)*	0.73 (0.96)	0.26 (0.34)	1.51 (1.45)
canada	–0.04 (–0.15)	*–0.88 (–3.17)*	*–0.98 (–3.93)*	–0.22 (–0.61)
denmark	0.40 (1.26)	0.34 (0.93)	0.20 (0.59)	0.41 (0.91)
spain	0.48 (1.00)	–0.07 (–0.12)	–0.30 (–0.57)	0.13 (0.23)
finland	0.94 (1.89)	0.35 (0.61)	0.36 (0.67)	–0.40 (–0.52)
france	0.34 (1.08)	0.03 (0.06)	–0.04 (–0.12)	0.59 (1.26)
greece	*2.78 (2.88)*	1.27 (1.14)	0.83 (0.71)	–0.57 (–0.34)
hong kong	0.94 (1.70)	0.07 (0.10)	0.48 (0.80)	*2.25 (2.69)*
indonisia	0.13 (0.14)	–1.76 (–1.64)	*–2.28 (–2.66)*	–1.10 (–0.98)
india	*1.97 (1.97)*	0.93 (0.83)	–0.06 (–0.05)	2.92 (1.89)
ireland	*0.66 (2.12)*	0.36 (0.98)	0.05 (0.14)	0.23 (0.49)
italy	0.26 (0.67)	–0.37 (–0.80)	–0.70 (–1.59)	–0.03 (0.00)
japan	–0.07 (–0.19)	*1.20 (3.13)*	*1.09 (3.02)*	–0.06 (–0.10)
korea	0.60 (0.87)	–0.89 (–1.28)	–0.21 (–0.33)	*–1.91 (–2.28)*
luxemburg	0.61 (1.79)	0.29 (0.72)	0.00 (–0.01)	0.36 (0.70)

	Standard one-factor	Standard four-factor	Standard InCapm	Alternative InCapm
mexico	1.40 (1.84)	1.05 (1.17)	0.67 (0.88)	0.70 (0.68)
malaysia	0.63 (1.06)	0.50 (0.74)	1.06 (1.75)	*2.22 (2.62)*
nederland	0.30 (1.10)	–0.02 (–0.06)	–0.05 (–0.17)	0.48 (1.24)
norway	0.40 (0.97)	–0.16 (–0.34)	–0.14 (–0.30)	0.31 (0.50)
new zeeland	0.67 (1.26)	–0.08 (–0.12)	–0.17 (–0.29)	0.15 (0.20)
austria	0.24 (0.67)	0.18 (0.41)	0.20 (0.49)	0.41 (0.76)
peru	*2.04 (2.05)*	1.86 (1.50)	1.99 (1.51)	3.20 (1.92)
philipinnes	1.46 (1.91)	1.00 (1.16)	1.15 (1.43)	1.74 (1.52)
portugal	0.03 (0.07)	–0.09 (–0.17)	–0.47 (–0.91)	0.80 (1.16)
south africa	0.33 (0.65)	0.13 (0.22)	0.47 (0.92)	1.42 (1.96)
sweden	*0.81 (2.07)*	0.00 (0.00)	0.11 (0.24)	0.47 (0.77)
singapore	0.24 (0.56)	–0.08 (–0.16)	0.16 (0.36)	0.33 (0.53)
switzerland	–0.11 (–0.47)	–0.38 (–1.40)	–0.43 (–1.93)	0.06 (0.21)
taiwan	1.01 (0.98)	0.99 (0.84)	*2.29 (2.07)*	1.08 (0.69)
thailand	0.25 (0.37)	–0.34 (–0.46)	0.12 (0.17)	0.35 (0.37)
uk	*0.53 (2.05)*	–0.09 (–0.31)	–0.24 (–0.94)	–0.15 (–0.41)
us	0.31 (1.25)	*–0.63 (–2.93)*	*–0.63 (–3.19)*	–0.27 (–0.87)
# Sig	7	3	6	4
Adj R^2	0.19	0.22	0.33	0.33
χ^2-test	0.04 (55.47)	0.008 (63.34)	0.00 (81.31)	0.004 (66.31)

Italic signals significance at a 5% level using GMM standard error taking into account cross-equation correlation and intertemporal hetero-scedasticity; # Sig is the number of significant alphas; Adj R^2 is the average adjusted R-squared; and χ^2-test is the p-value of the Wald test (H_0: all alphas equal to zero).

breakpoints are calculated (only NYSE stocks, or all stocks, including Amex and NASDAQ ones); and (*vii*) the correction for intertemporal hetero-skedasticity (OLS v GMM). We also found that size-bias in the database seems to have a large effect, especially on the size-factor portfolio and to some extent we discovered interaction effects between the influences of the three well-known anomalies on expected stock returns.

Elements that did not have a large influence were: (*i*) the time period (80-93, 94-2000 or 80-2000); (*ii*) risk-free rate and market rate (IMF average 3-month US T-bill or IMF end-of-period discount rate; DataStream's US-market index or computing the market return as the value-weighted return on all stocks in the size-distress portfolios plus the negative book value equities as in Fama and French (1993)); (*iii*) the way of calculating the breakpoints (broad- or narrow-based breakpoints i.e. are stocks with unknown market value or book-to-market value included or not).

We also found that both survivorship- and size-bias do not have a substantial influence on the distress- and momentum factor portfolios, and the size-factor portfolio is not influenced much by survivorship-bias. We draw special attention to the possible underestimation of the small firm effect by value-weighting the portfolio returns, calculating the decile breakpoints on NYSE stocks only and excluding stocks with an unknown market value or book-to-market value from the database.

The way of composing the right-side factor portfolios also influences the performance of an asset pricing model. We propose an alternative way of handling the three standard factors (size, distress and momentum) that produces estimated alphas closer to zero. Notably, the size and distress factor portfolios are split up into two factor portfolios, one for the smallest or most distressed stocks risks and one for regular size and distress risks. This alternative model significantly improves on the standard four-factor CAPM model nesting Fama and French ((1993), (1995), (1996a), (1996b)), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999), and this superior performance is illustrated for one-dimensional size-, distress-, momentum- and industry portfolios and two-dimensional size-distress portfolios of US stocks. To reduce the risk of *ad hoc* modifications, we also test the proposed modification in an international setting of 39 countries, both emerging and developed. We showed that the factors that performed well in the US-market also work in an international setting.

VI. FURTHER RESEARCH

We did not test, at this stage, the ability of the standard four-factor model and the "alternative" model to price other two-dimensional test-portfolios (*e.g.* size-momentum and distress-momentum) nor three dimensional test-portfolios (*e.g.* size-distress-momentum). The specific choice of the exchange factor portfolios in the international asset pricing models remains arbitrary and is subject to further research, as is the significance of the exchange-rate factors as a group. We could extend the international model even more by allowing two versions of each size-, B/M- and momentum factor portfolio: one composed out of the "emerging" stock basket and the other out of "developed" stock basket; moreover, each version can have its own specific factor portfolio composition.

NOTES

1. Fama and French (1993) do use only the NYSE stocks to set allocation breakpoints for both size and distress, not the median of all stocks (including Amex and NASDAQ). The reason for this is not stated explicitly.
2. With identical regressors across equations and no cross-equation restrictions, Seemingly Unrelated Regression (SUR) provides the same estimates and standard errors as OLS. GMM, with as constraint zero-mean residuals orthogonal on the factor returns within each time series, also collapses to OLS. But the weighting matrix we use is White's hetero-skedasticity-consistent covariance matrix, so that the significance statements are robust to both hetero-skedasticity (over time or across stocks) and contemporaneous correlation of unknown form.
3. Fama and French (1993) exclude financial firms because the high leverage that is normal for these firms probably does not have the same meaning as for non-financial firms, where high leverage more likely indicates distress.
4. For reasons of availability and international comparability our risk-free rates are from the IMF's *International Financial Statistics*. For the US T-bill, this source provides only a monthly average of the 3-month rate. The IMF US discount rate, in contrast, is an end-of-period rate and is available for most countries. Eligible depository institutions pay this rate when borrowing short-term from a Federal Reserve Bank.
5. The fact S is actually computed from the (value-weighted) returns of three size/distress intersections, only partly mitigates this effect, because the relation between size and distress is far from perfect.
6. As an alternative, we tried working with unions rather than intersections. Under that procedure we allocate every stock with a known market value into one of five size-groups and all stocks with a known book-to-market into one of five distress-groups. This gives us ten basic test portfolios, We then form a portfolio for size/distress combination (*i*, *j*) as the average of size portfolio *i* and distress portfolio *j*, weighted by the number of stocks in *i* and *j*, respectively, thus computing 25 different combinations of the ten basic alphas, similar to the 25-portfolio tests used thus far. The outcome was a somewhat larger number of rejections (nine, up from seven) despite generally lower alphas – a signal of higher power relative to the original FF design but nowhere as

strong as the alternative procedure outlined in the main text. In addition, this induces strong dependencies across tests.

7. Besides allowing more attention to small stocks, another consideration for basing the breakpoints on the entire sample rather than the list of the leading exchange is that the latter procedure cannot be applied consistently across countries.
8. A possible explanation of that view is that the price behavior of large stocks on average depends on the transactions of less professional or less dedicated investors. Less dedicated investors are likely to transact in larger firms due to extended media attention towards larger firms and they are likely to keep away from smaller unknown firms. A less dedicated investor, however, has nor the means, the time or the skill to transact much or quickly (large transaction costs, slow information, less active trader). This could induce momentum in the larger stocks. Professional or dedicated investors invest more in smaller stocks compared to the less dedicated non-professional investor. So the price behavior of smaller stocks will on average depend on the transactions of more dedicated investors. However, a professional or dedicated investor has more means, time or skill to transact much and quickly (low transaction costs, fast information, more active trader). This could induce short term reversal into the small stock returns.
9. An alternative procedure might have been to find a transformation of the factors such that the risk-return relationship becomes linear.
10. Fama and French (1992) use the CRSP database
11. Eun, Huang and Lai (2003) use DataStream's Market Data
12. We are aware that only a Wald test on all alphas is the appropriate test for an asset pricing model. But we show the individual t-statistics to get a feeling of which portfolios might be priced badly and how we might make the model better.
13. Also in Table 18 our alternative momentum model does slightly better than the standard momentum model in the sense that the Wald statistic is lower.
14. Fama and French (1995) investigate size and book-to-market factors, a momentum factor is not included.
15. Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Denmark, Finland, France, Germany, Greece, Hong Kong, India, Indonesia, Ireland, Italy, Japan, Luxemburg, Malaysia, Mexico, Netherlands, New Zealand, Norway, Peru, Philippines, Portugal, Singapore, South Korea, South-Africa, Spain, Sweden, Switzerland, Taiwan, Thailand, United Kingdom, United States
16. Thus, *SMB* (small minus big) is the size factor portfolio: a zero-investment portfolio that is long the 50% smallest stocks and short the 50% largest stocks; *HML* (high minus low) is the distress factor portfolio: a zero-investment portfolio that is long the 30% highest B/M stocks and short the 30% lowest B/M stocks; and *WML* (winner minus loser) is the momentum factor portfolio: a zero-investment portfolio that is long the 30% highest past-performers (winners) and short the 30% lowest past-performers (losers).
17. In a nutshell, an international capital asset pricing model (InCAPM) takes into account possible real exchange rate risks because every investor measures returns in the real terms that are specific for her country rather than in a common currency (the USD, here). However, inflation differentials are dwarfed by exchange-rate changes, so that the real rate closely tracks the nominal one; and inflation rates are virtually uncorrelated with stock returns anyway. Thus, following standard practice we use nominal exchange-rate changes rath

REFERENCES

Adler, M. and B. Dumas, 1984, Exposure to Currency Risk: Definition and Measurement, *Financial Management* 13, 41-50.

Allayannis, G., 1995, Exchange Rate Exposure Revisited, 55th Annual American Finance Association Conference, (Washington), January 6-8.

Allayannis, G., 1997, The Time-Variation of the Exchange Rate Exposure: an Industry Analysis, 57th Annual American Finance Association Conference, (New Orleans), January 5.
Banz, Rolf W., 1981, The Relationship between Return and Market Value of Common Stocks, *Journal of Financial Economics* 9, 3-18.
Bartov, E. and G.M. Bodnar, 1994, Firm Valuation, Earnings Expectations and the Exchange-Rate Exposure Effect, *Journal of Finance* 49, 1755-1785.
Bodnar, G. and W. Gentry, 1993, Exchange-Rate Exposure and Industry Characteristics: Evidence from Canada, Japan and U.S., *Journal of International Money and Finance* 12, 29-45.
Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57-82.
Eun C., Huang W. and S. Lai, 2003, International Diversification with Large- and Small-Cap Stocks, Working paper.
Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427-465.
Fama, Eugene F., and Kenneth R. French, 1993, Common Factor Portfolios in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3-56.
Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance* 50, 131-155.
Fama, Eugene F., and Kenneth R. French, 1996a, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance* 51, 55-84.
Fama, Eugene F., and Kenneth R. French, 1996b, The CAPM is Wanted, Dead or Alive, *Journal of Finance* 51, 1947-1958.
Fama, Eugene F., and Kenneth R. French, 1998, Value Versus Growth: the International Evidence, *Journal of Finance* 53, 1975-1999.
Fama, Eugene F., and Kenneth R. French, 2000, Characteristics, Covariances, and Average Returns: 1929 to 1997, *Journal of Finance* 55, 389-406.
Jorion, P., 1990, The Exchange Rate Exposure of U.S. Multinationals, *Journal of Business* 63, 331-345.
Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* 48, 65-91.
Rees, William and Sanjay Unni, 1999, Exchange Rate Exposure amongst European Firms: Evidence from France, Germany and the UK, (University of Glasgow), AccFin Research Paper 1999/8
Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Pervasive Evidence of Market Inefficiency, *Journal of Portfolio Management* 9, 18-28.
Rouwenhorst, Geert K., 1999, Local Return Factors and Turnover in Emerging Stock Markets, *Journal of Finance* 54, 1439-1464.
Sercu, Piet and Raman Uppal, 1995, International Financial Markets and the Firm, (South-Western College Publishing), 347-367, 583-619.
Sercu, Piet, 1980, A Generalization of the International Asset Pricing Model, *Revue de l'Association Française de Finance* 1, 1, 91-135.
Solnik, H. Bruno, 1974a, An Equilibrium Model of the International Capital Market, *Journal of Economic Theory* 8, 500-524.
Solnik, H. Bruno, 1974b, The International Pricing of Risk: an Empirical Investigation of the World Capital Market Structure, *Journal of Finance* 29, 365-378.
Solnik, H. Bruno, 1977, Testing International Asset Pricing: some Pessimistic Views, *Journal of Finance* 32, 503-512.
Solnik, H. Bruno, 1983, International Arbitrage Pricing Theory, *Journal of Finance* 38, 449-457.
Stattman, Dennis, 1980, Book Values and Stock Returns, *The Chicago MBA: a Journal of Selected Papers* 4, 25-45.